

TARTU RIIKLIKU ÜLIKOOLI  
**TOIMETISED**

УЧЕНЫЕ ЗАПИСКИ  
ТАРТУСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА  
ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS

551

ЛОГИКО-СЕМАНТИЧЕСКИЕ  
ВОПРОСЫ ИСКУССТВЕННОГО  
ИНТЕЛЛЕКТА

Труды по искусственному интеллекту

III

TARTU RIIKLIKU ÜLIKOOLI TOIMETISED  
УЧЕНЫЕ ЗАПИСКИ  
ТАРТУСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА  
ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS  
ALUSTATUD 1893.a. VIHK 551 ВЫПУСК ОСНОВАНЫ В 1893.g.

# ЛОГИКО-СЕМАНТИЧЕСКИЕ ВОПРОСЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Труды по искусственному интеллекту

III

ТАРТУ 1980

Р е д а к ц и о н н а я   к о л л е г и я : И.Куль,  
Х.Рятсеп, И.Сильдмяэ, Ю.Тулдава, К.Ээремаа

Ответственный редактор выпуска: Х.Ыйм

# СОДЕРЖАНИЕ - CONTENTS

<b>Д.Д. Апресян. О поверхностно-семантическом компоненте модели "Смысл <math>\longleftrightarrow</math> Текст" .....</b>	<b>5</b>
<b>Ju. Apresjan. On the surface semantic component in the "Meaning-Text" model. Summary .....</b>	<b>27</b>
<b>В.А. Звегинцев. К стратегии построения систем класса "искусственный интеллект" .....</b>	<b>28</b>
<b>V. Zvegintsev. Some notes on the strategy of building systems of the "Artificial Intelligence" class. Summary.....</b>	<b>41</b>
<b>О.В. Звегинцева. Построение редуцированного языка для автоматических диалоговых систем .....</b>	<b>43</b>
<b>O. Zvegintseva. Creating a reduced language for automatic dialog systems. Summary.....</b>	<b>50</b>
<b>Л.Д. Йомдин. К построению формальной модели современного русского языка. Оператор синтаксического согласования для определительных конструкций ....</b>	<b>51</b>
<b>L. Yomdin. On working out a formal model of modern Russian: an operator of syntactic agreement for attributive constructions. Summary.....</b>	<b>70</b>
<b>М. Койт. Алгоритмы перевода с <math>\psi</math>-языка на русский ....</b>	<b>71</b>
<b>M. Koit. Two algorithms for translating from <math>\psi</math>-language into Russian. Summary.....</b>	<b>81</b>
<b>С.Р. Литвак, Т.А. Роосмаа, М.Э. Салувер, Х.Я. Ним. Об автоматическом морфологическом анализе ограниченного естественного языка .....</b>	<b>82</b>
<b>S. Litvak, T. Roosmaa, M. Saluveer, H. Õim. Automatic morphological analysis of limited natural language. Summary.....</b>	<b>86</b>
<b>С.Р. Литвак, Т.А. Роосмаа, М.Э. Салувер, Х.Я. Ним. Подсистема автоматического синтаксического анализа для экспериментальной ВОС .....</b>	<b>87</b>
<b>S. Litvak, T. Roosmaa, M. Saluveer, H. Õim. A subsystem of automatic syntactic analysis for an experimental question-answering system. Summary.....</b>	<b>91</b>

<b>Д.А. Поспелов, В.А. Шустер. Нормативное поведение интеллектуальных систем .....</b>	<b>92</b>
D. Pospelov, V. Shuster. Normative behavior of intellectual systems. Summary.....	108
<b>В.М. Сафонов. Проблемы построения автоматизированных имитаторов процесса творчества .....</b>	<b>109</b>
V. Safonov. Some problems connected with the construction of automatic imitators of creative processes. Summary.....	120
<b>И. Сильдмэе. Значение, смысл и мысль .....</b>	<b>121</b>
I. Sildmäe. Meaning, sense, thought. Summary.....	128
<b>Д.И. Шапиро, В.Н. Васильев. Диалоговая система принятия решений с контролем состояния ДНР .....</b>	<b>129</b>
D. Shapiro, V. Vasilyev. A dialog system for decision making with following check-up. Summary..	135

# О ПОВЕРХНОСТНО-СЕМАНТИЧЕСКОМ КОМПОНЕНТЕ В МОДЕЛИ "СМЫСЛ $\longleftrightarrow$ ТЕКСТ"

Ю. Д. Апресян

## 1. Общая характеристика модели "Смысл $\longleftrightarrow$ Текст"

В работе принимается и предполагается известной читателю общая концепция модели "Смысл  $\longleftrightarrow$  Текст", предложенной в [12] (см. также [8]).

Модель "Смысл  $\longleftrightarrow$  Текст" определенного естественного языка представляет собой единое формальное описание его грамматики и словаря, реализованное в виде системы правил, которые функционально подобны носителям этого языка.

Концепция "единого" или "интегрального" лингвистического описания, известная в лингвистике по крайней мере с 1964 года [24] и представляющая собой одно из главных новшеств современной теоретической мысли, применительно к модели "Смысл  $\longleftrightarrow$  Текст" может быть представлена следующими тезисами:

(I i) Словарь является неотъемлемой частью полного теоретического описания языка, во всех существенных отношениях уравненной с грамматикой;

(I ii) Формальный словарь и формальная грамматика как части единой теории языка должны быть идеальным образом согласованы друг с другом в отношении распределения языковой информации, форм ее записи и взаимодействия грамматических правил со словарными материалами (правилами и объектами)<sup>1</sup>;

(I iii) Требование согласования грамматической и словарной информации включает в себя, в частности, требование того, чтобы значения всех содержательных единиц языка - слов, фразем, граммем, конструкций, просодических элементов - были описаны на одном и том же семантическом языке (ср. [22])<sup>2</sup>.

С другой точки зрения модель "Смысл  $\longleftrightarrow$  Текст" представляет собой многоуровневый транслятор смыслов в предложения и наоборот. Выделяются четыре основных уровня - фонетический, морфологический, синтаксический и семантический. Каждый из них, за исключением семантического, подразделяется на два других уровня - поверхностный и глубокий; нам аналогичное деление представляется необходимым и на семантическом уровне (см. об этом ниже).

На каждом из восьми уровней исходное предложение (при анализе) или исходный смысл (при синтезе) получает определенный формальный образ, называемый представлением предложения соответствующего уровня. Мы будем иметь дело главным образом с глубинно-синтаксическими и поверхностно-семантическими представлениями.

Представление предложения на каждом из двух интересующих нас уровней состоит из а) структуры соответствующего уровня, б) сведений о коммуникативной организации предложения (тема и рема, данное и новое<sup>3</sup>, пресуппозиция и утверждение, логические акценты и эмфатические выделения), в) сведения о референтности - нереферентности - кореферентности различных узлов, г) указаний об областях действия одних единиц относительно других.

Главной частью представления предложения является его структура. Специальный интерес в рамках данной работы представляют два типа структур - глубинно-синтаксическая и поверхностно-семантическая.

## 2. Глубинно-синтаксическая структура (ГСС).

ГСС - это дерево зависимостей с составляющими, т.е. такое дерево, в котором стрелки зависимостей могут соединять не только узлы, но и группы узлов<sup>4</sup>. В узлах ГСС стоят имена лексем глубинно-синтаксического языка с необходимыми характеристиками, а стрелки зависимостей соответствуют одному из универсальных (единых для всех языков) подчинительных отношений глубинного синтаксиса. Охарактеризуем сначала словарь глубинно-синтаксического языка, а затем - его синтаксис.

К числу словарных единиц глубинно-синтаксического языка относятся а) все семантически содержательные слова естественного языка, б) фразеологические единицы, каждая из которых стянута в единый узел, в) фиктивные глубинно-синтаксические слова, являющиеся интерпретацией семантически содержательных поверхностно-синтаксических отношений, г) фиктивные потенциальные слова, предсказываемые словообразовательной системой языка в целом и вводимые для сохранения общности некоторых правил перифразирования.

В число глубинно-синтаксических отношений входят 6 актантных отношений (считается, что более чем шестиместных предикатов в естественных языках нет, - см [1]), а также координативное, атрибутивное и присоединительное отношения. Координативное отношение описывает разного рода сочинитель-

ные конструкции, атрибутивное - определительные, обстоятельственные, аппозитивные, композитные, количественные, ограничительные, разъяснительные, собственно атрибутивные и некоторые другие конструкции, присоединительное - разного рода вводные конструкции.

В связи с тем, что синтаксис глубинного языка является, по замыслу, универсальным, глубинно-синтаксические отношения должны быть асемантичны. Это значит, что все семантически содержательные поверхностно-синтаксические отношения на глубинно-синтаксическом уровне должны быть проинтерпретированы, т.е. переработаны в фиктивные слова группы в). Таковы, в частности, следующие поверхностно-синтаксические отношения: модально-инфинитивное (2 i), субъектно-темпоральное (2 ii), порционно-темпоральное (2 iii), аппроксимативно-количественное (2 iv), условно-предикативное (2 v) - см. [12, 232].

(2 i) Быть грозе  $\cong$  Должна быть гроза,

(2 ii) Молодым он ходил всю зиму без шапки  $\cong$  Когда он был молодым, он ходил всю зиму без шапки,

(2 iii) Больной часами не выходил из комнаты  $\cong$  Обычно больной не выходил из комнаты очень долго - по несколько часов,

(2 iv) Ему было лет тридцать  $\cong$  Ему было приблизительно тридцать лет,

(2 v) Останься он дома, все было бы в порядке  $\cong$  Если бы он остался дома, все было бы в порядке.

В предложениях (2 i) - (2 v) существенная часть смысла, которую мы постарались отразить в даваемых после знака прилизительности перифразах, выражается очень неявно - падежами и числом существительных и прилагательных, личными и неличными формами глагола, порядком слов. На поверхностно-синтаксическом уровне соответствующие компоненты смысла предложения удерживаются в названиях поверхностно-синтаксических отношений, а наиболее естественный способ сохранить их и на глубинно-синтаксическом уровне состоит в том, чтобы ввести фиктивные глубинно-синтаксические слова, например, слово ДОЛЖЕН для (2 i), слово ОДНОВРЕМЕННОСТЬ - для (2 ii), слово ПРИБЛИЗИТЕЛЬНО - для (2 iv) и т.д.

Отметим, что каждое фиктивное глубинно-синтаксическое слово должно иметь самостоятельное толкование, отличное от толкования внешне совпадающего с ним слова естественного языка. В словаре у фиктивного слова должен быть особый вход.



Из слов и отношений глубинно-синтаксического языка строятся ГСС. В узлах ГСС стоят имена лексем глубинно-синтаксического языка с необходимыми характеристиками. В число необходимых характеристик узла, помеченного именем нефиктивной лексемы, входят сведения о его глубинно-синтаксических признаках, т.е. о том, значении какой лексической функции от какого аргумента он является, и сведения о семантически содержательных граммемах, т.е. тех граммемах, выражение которых диктуется существом сообщаемого, а не согласовательными синтаксическими правилами языка.

ГСС предложения

(31) Похолодание вызвало смерть черепах выглядит (с некоторыми упрощениями) следующим образом:

(31)



Из (31) по весьма общим лексическим и синтаксическим правилам перифразирования, определенным в глубинно-синтаксическом языке, могут быть получены ГСС предложений (311) - (3vii) и многих других, синонимичные (31):

(311) Похолодание вызвало гибель черепах (замена смерть  $\Leftrightarrow$  гибель по правилу  $X \Leftrightarrow \text{Syn}(X)$ ),

(31ii) Похолодание привело к смерти черепах (замена вызвать  $\Leftrightarrow$  привести к по тому же правилу),

(3iv) Смерть черепах проистекла из-за похолодания (замена вызвать  $\Leftrightarrow$  проистечь из-за по правилу  $X \Leftrightarrow \text{Cov}_{21}(X)$ ),

(3v) Похолодание было причиной смерти черепах (замена вызвать  $\Leftrightarrow$  быть причиной по правилу  $X \Leftrightarrow \text{Cov}_{11} \xrightarrow{2} s1(X)$ ),

(3vi) Смерть черепах была следствием похолодания (замена вызвать  $\Leftrightarrow$  быть следствием по правилу  $X \Leftrightarrow \text{Cov}_{11} \xrightarrow{2} s2(X)$ ),

(3 vii) Черепашки умерли из-за похолодания (замена вызвать смерть  $\longleftrightarrow$  умереть из-за по правилу  $X \xrightarrow{2} Y \longleftrightarrow Y \xrightarrow{attr} Adv1B(X)$ ).

### 3. Поверхностно-семантическая структура (ПСемС).

В первоначальной версии модели "Смысл  $\longleftrightarrow$  Текст" не предполагалось, что семантический уровень представления предложений подразделяется на подуровни. Основным объектом этого уровня было семантическое представление, а его главным компонентом - семантическая структура (СемС), которая имеет вид связанного ориентированного графа. Каждый из узлов графа соответствует некоему семантическому конструкту - слову семантического языка. Разумно допустить, что это может быть граф с "мешками", т.е. такой граф, ребра (дуги) которого соединяют не только отдельные узлы, но и группы узлов.

Вся чисто семантическая информация, содержащаяся в СемС, сосредоточена исключительно в узлах графа. Дуги графа - стрелки зависимостей - сами по себе не несут никакой смысловой информации, а указывают лишь на синтаксическую организацию смыслов (узлов)<sup>5</sup>.

Одному и тому же предложению может соответствовать несколько СемС, но эти СемС отличаются друг от друга не принципиально, а лишь степенью детальности представления смысла. Среди всех СемС имеется одна каноническая, составленная исключительно из простейших (неопределяемых) понятий, своего рода смысловых атомов, различные комбинации которых формируют значения реальных единиц естественных языков.

Язык семантических графов является, по предположению, универсальным - единым для всех человеческих языков. Очевидно, что тогда все синонимичные друг другу предложения всех естественных языков должны быть отображены в одно и то же семантическое представление (СемП) - их единый универсальный образ на семантическом языке.

Создание такого семантического языка продолжает оставаться конечной задачей лингвистической семантики, но перспектива получения надежных результатов в этой области, если исходить из реального опыта современных исследований, несколько отдалается.

Прежде всего следует подчеркнуть, что никто из современных теоретиков семантики, за исключением, может быть, А.Важбицкой [30], не работает со столь глубокими СемС. Большинство исследователей используют гораздо более поверхностный

семантический язык (ср. [21], [12], [15]), а многие вообще не прибегают ни к какому специальному языку, довольствуясь перифразированием в качестве основного средства экспликации смысла предложений (ср. [27], [26], [19]).

Нам кажется, что это обстоятельство не случайно, что оно **отражает** не каприз исследователей и не переходящий этап в истории развития науки, а некоторые фундаментальные свойства объекта изучения.

Главным из этих свойств мы считаем национальное своеобразие семантики естественного языка. Это своеобразие имеет много различных проявлений, из которых в первую очередь следует упомянуть факт обязательности выражения некоторых смыслов в предложении.

Обычно, когда говорят об обязательности выражения некоторых смыслов в определенном естественном языке, ссылаются на значения грамматических категорий ([18], [23], [11]). Хрестоматийными примерами такого рода являются вид в славянских языках и артикль в романских. Русская глагольная форма, например, через грамматическую категорию вида выражает идею законченности или незаконченности действия, а французская глагольная форма, в особенности **такая**, как инфинитив или причастие, может оставлять эту идею невыраженной<sup>6</sup>. С другой стороны, при французском существительном через артикль обязательно выражается значение определенности - неопределенности, а при русском существительном оно выражается лишь тогда, когда это диктуется содержанием сообщения.

Следует сказать, что обязательность выражения определенных значений свойственна не только грамматике, но и лексике. Можно, например, заметить, что русский язык по сравнению с французским навязывает говорящим гораздо более детальное описание действий с точки зрения способа их осуществления, независимо от того, требует ли существо сообщаемого такой детализации или нет.

Смысл 'перемещаясь, перестать находиться где-либо' выражается в русском языке разными глаголами - выйти, вылететь, выплыть, выползти - в зависимости от того, как перемещается субъект (идет, летит, плывет или ползет), между тем как во французском языке тот же смысл выражается единообразно - глаголом *sortir*. Хотя в русском языке имеется глагол покинуть, являющийся родовым обозначением соответствующего действия, в стилистически и семантически нейтральном контексте

следует говорить (41), а не (411):

(41) Собака вышла из конуры,

(411) ?Собака покинула конуру.

С другой стороны, по-французски можно сказать *sortir en marchant*, если нужно в подчеркнутой форме выразить ту мысль, что некое существо, покинувшее какое-то место, именно шло, а не перемещалось иным способом, но в условиях, когда такого специального смыслового задания нет, следует употреблять просто глагол *sortir*. В нейтральном контексте говорят (51), а не (511):

(51) *Le chien est sorti de sa niche,*

(511) ?*Le chien est sorti en marchant de sa niche.*

Подобные расхождения свойственны не только глаголам со значением перемещения, но и глаголам со значением положения в пространстве. По-русски несколько необычны или сомнительны фразы:

(61) Поль находился у окна,

(71) ?Не находишься у окна!

Соответствующие мысли должны быть выражены фразами

(611) Поль стоял у окна,

(711) Не стой у окна!

хотя для говорящего может быть несущественным тот факт, что субъект этих ситуаций находился у окна в вертикальном положении, опираясь на ноги. Этот "лишний" смысл выражается в русском языке автоматически. Адекватные французские переводы фраз (611) и (711) - фразы (81) и (91) - по своему буквальному смыслу больше соответствуют фразами (61) и (71):

(81) *Paul se trouvait près de la fenêtre,*

(91) *Ne reste pas près de la fenêtre!*

Выделенный разрядкой смысл выражается во французском языке лишь в том случае, когда в этом есть содержательная необходимость:

(811) *Paul se trouvait debout près de la fenêtre,*

(911) *Ne reste pas debout près de la fenêtre!*

Здесь обнаруживается глубокая аналогия между морфологическими и семантическими фактами. Известно деление морфологических категорий на синтаксически обусловленные и необусловленные (ср. синтаксическую обусловленность числа, рода, лица у глагола, падежа, числа, рода у прилагательного, падежа у существительного и синтаксическую необусловленность

времени, вида, **наклонения** у глагола, степени сравнения у прилагательного, числа у существительного). Известно, с другой стороны (см. [10]), что почти для каждой синтаксически обусловленной категории существует синтаксическая конструкция, в которой она оказывается семантически содержательной (ср., например, число прилагательного в сочинительных конструкциях типа красный и белый шары, где значение **один красный шар и один белый шар** выражается именно числом прилагательного, а не существительного).

В поверхностно-синтаксических структурах сведения о синтаксически обусловленных категориях **элиминируются**, за исключением тех случаев, когда последние семантически содержательны.

Рассмотренные выше факты показывают, что с автоматически выражаемыми значениями, независимо от их природы, дело обстоит до известной степени **похожим** образом: в одних случаях выражение такого значения может быть вынужденным, т.е. чисто автоматическим, а в других (например, в условиях просодической выделенности и противопоставления) — полностью совпадать с намерениями говорящего.

Очевидно, что правила обращения с такими значениями должны быть аналогичны правилам обращения с синтаксически обусловленными категориями: в ситуациях первого рода автоматически выражаемые значения **нерелевантны и могут** быть сняты (или порождены) без ущерба для существа сообщения, а в ситуациях второго рода они релевантны для сообщения и должны **оставаться неизменными**.

Таким образом, в смысле предложения естественного языка мы обнаруживаем два разных слоя, или уровня — национальный и универсальный. Чтобы добраться до универсального, самого глубокого слоя семантики, необходимо в определенных ситуациях уметь снимать обязательно выражаемые в данном языке, т.е. национально своеобразные значения.

Уже из этих соображений следует, что в формальной модели "Смысл  $\longleftrightarrow$  Текст" должно быть два уровня представления смысла — поверхностно-семантический (уровень национальной семантики) и глубинно-семантический (уровень универсальной семантики). На первом из них отражаются все выраженные в соответствующем предложении значения, в том числе и нерелевантные, а на втором — только те, которые существенны для сообщения.

Имеется еще одно основание для расширения семантическо-

го уровня на поверхностно- и глубинно-семантический подуровни.

Как мы уже говорили, модель "Смысл  $\longleftrightarrow$  Текст" для каждого естественного языка должна формально эксплицировать отношения синонимии между предложениями этого языка. В смысле синонимичных предложений могут входить как раз те обязательные, национально своеобразные значения, о которых мы говорили выше. Следовательно, синонимические отношения между ними должны устанавливаться раньше, чем на глубинно-семантическом уровне, где по крайней мере часть этих значений может оказаться снятой.

Естественно возникает вопрос, могут ли они все же описываться не на поверхностно-семантическом, а на глубинно-синтаксическом уровне, где действует специально для этого предназначенная система правил перифразирования (см. (31) - (3vii)). Оказывается, однако, что на глубинно-синтаксическом уровне эксплицируются далеко не все отношения такого рода. Синонимическое перифразирование (31) - (3vii) обеспечивается глубинно-синтаксическими правилами, а синонимическое перифразирование (31)  $\longleftrightarrow$  (10) - нет:

(31) Похолодание вызвало смерть черепах,

(10) Понижение температуры воздуха было причиной смерти черепах.

Синонимия (31) - (10) устанавливается не непосредственно, а через поверхностно-семантический уровень, на котором обе ГСС отражаются одинаково в виде ПСемС (3'), приводимой ниже.

Семантическими разложениями содержательных единиц языка, относящимися не более чем к поверхностно-семантическому уровню, необходимо располагать и для экспликации отношений синонимии между предложениями (некоторые замечания на эту тему см. в [3, 19-20]).

Итак, на поверхностно-семантическом уровне - уровне национальной семантики - решаются следующие задачи: а) эксплицируются все выраженные в предложении значения, в том числе и те обязательные (обусловленные) значения, которые семантически нерелевантны; б) эксплицируются отношения синонимии между предложениями данного естественного языка ("национальная" синонимия) и тем самым показываются системные семантические связи между различными содержательными единицами языка, в том числе и единицами разной природы; в) эксплицируются

отношения омонимии между предложениями данного языка.

На глубинно-семантическом уровне - уровне универсальной семантики - решаются следующие задачи: а) элиминируются все семантически нерелевантные значения; б) эксплицируются отношения синонимии между предложениями разных (в идеале - всех) естественных языков.

Предлагаемое нами расщепление прежде единого семантического уровня на два новых подуровня обнаруживает неслучайное сходство с соответствующим делением синтаксического уровня: поверхностно-синтаксический уровень - это уровень, на котором описываются специфические для данного языка "национальные" синтаксические конструкции, а глубинно-синтаксический уровень - это уровень, на котором элиминируется национальная синтаксическая специфика, т.е. уровень универсального синтаксиса, единого для всех естественных языков.

Теперь попытаемся представить в самых общих чертах языки обоих уровней представления смысла предложений и семантические структуры, которые строятся из материала этих языков.

В свете того, что было сказано выше, глубинно-семантическим языком разумно считать элементарный универсальный семантический язык, словарь которого состоит исключительно из неопределяемых понятий - семантических конструкторов, являющихся именами предметов или именами простейших предикатов, по преимуществу одно- или двухместных. Из таких слов по правилам предикатно-аргументного синтаксиса строятся глубинно-семантические структуры - по существу, те канонические структуры, о которых шла речь выше.

Поверхностно-семантический язык отличается от глубинно-семантического и своим словарем, и своим синтаксисом.

Поскольку поверхностно-семантический язык предназначен для экспликации национальной семантики, основная часть его словаря должна формироваться из наименее идиоматичных, достаточно простых, но не обязательно элементарных слов того же естественного языка. Они должны быть отобраны таким образом, чтобы внутри поверхностно-семантического языка они не имели синонимов и использовались каждое ровно в одном значении. Таким образом, словарь поверхностно-семантического языка представляет собой продукт радикального упрощения (сокращения не меньше чем на два порядка) и стандартизации (снятие синонимии, полисемии и омонимии) словаря соответствующего естественного языка.

С другой стороны, поскольку слова поверхностно-семантического языка в общем случае неэлементарны, его словарь значительно больше по объему, чем словарь универсального семантического языка, и помимо одно- и двухместных предикатов включает предикаты с большим числом мест.

Что касается различий поверхностно-семантического и глубинно-семантического языков в области синтаксиса, то они прямо вытекают из природы их предикатных слов: в число поверхностно-семантических отношений входят не два актантных отношения, как в глубинно-семантическом языке, а 5-6 актантных отношений. Синтаксис поверхностно-семантического языка, в отличие от его словаря, не извлекается из естественного языка, а от начала и до конца конструируется.

Из слов поверхностно-семантического языка и его синтаксических отношений строятся ПСемС. В достаточно большом числе случаев (но, естественно, далеко не всегда) ПСемС является продуктом первого шага замены содержательных единиц соответствующей ГСС (слов и граммем) их семантическими представлениями. Обязательной замене на семантические представления подлежат все граммеы и те из лексических единиц, которые не совпадают ни с одним словом поверхностно-семантического языка или ведут себя нестандартно относительно правил семантических преобразований.

В качестве иллюстрации приведем (с упрощениями) ПСемС (3') для ГСС (31); в нее отображается и предложение (10)<sup>8</sup>. Эта ПСемС отражает следующие содержательные представления о значениях использованных в (61) единиц.

Похолодание = уменьшение температуры атмосферного воздуха .

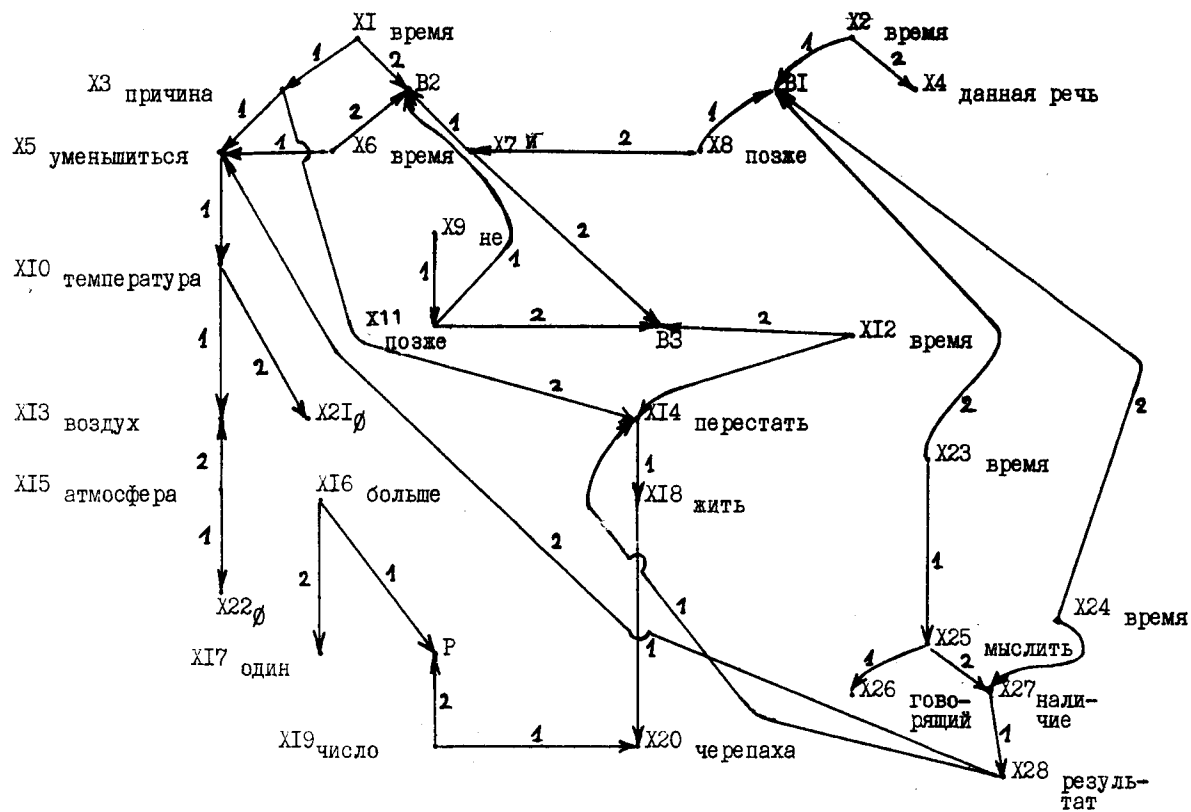
Х вызвал У = Во время  $T_2$  имела место ситуация Х; во время  $T_3$  имела место ситуация У; Х является причиной У-а;  $T_1$  - время данной речи - позже  $T_2$  и  $T_3$ ;  $T_2$  не позже  $T_3$ ; говорящий мыслит наличие результата в момент речи (несколько упрощенное толкование лексического значения глагола вызывать и граммем прошедшего времени, изъявительного наклонения и совершенного вида).

Смерть Х-ов = Тот факт, что Х, число которых больше одного, перестал жить (толкование лексического значения слова смерть и граммем множественного числа при Х).

При нынешнем состоянии разработки семантических проблем ничего более определенного о семантических языках, ГСемС и



(3')



ПСемС сказать нельзя. В связи с неясностью многих вопросов, касающихся, в частности, ПСемС, мы в данной работе пользуемся более привычной формой экспликации смысла в виде толкований, однако наши толкования (см. приведенные выше примеры) отличаются рядом существенных свойств и от перифраз, и от традиционных лексикографических дефиниций.

Под толкованием единицы X мы понимаем предложение на ограниченном и стандартизованном языке, построенное с соблюдением следующих требований: (а) оно должно быть в точности синонимично толкуемой единице X; (б) оно должно сводить толкуемый смысл к более простым смыслам - словам поверхностно-семантического языка (см. выше); (в) эти более простые смыслы должны быть настолько крупными, насколько это возможно без нарушения требования (б). Последнее условие значит, что толкуемую единицу X следует сводить к ее "непосредственным семантическим составляющим", а не к мельчайшим неопределяемым словам - смыслам. Обоснование этих требований содержится в [1].

Мы допускаем, что для решения трех основных задач поверхностно-семантического уровня - экспликации национальной семантики, а также отношений синонимии и омонимии между предложениями естественного языка - достаточно именно той степени детальности представления смысла, которая достигается в толкованиях языковых единиц. С другой стороны, такая степень детальности представления смысла абсолютно необходима: если бы мы не располагали толкованиями языковых единиц, а работали бы с самими этими единицами в собранном виде, мы бы лишились возможности эксплицировать отношения синонимии и омонимии между многими предложениями естественного языка с помощью простых и принципиальных правил.

4. Задача поверхностно-семантического компонента модели языка.

Задачей поверхностно-семантического компонента модели "Смысл  $\Leftrightarrow$  Текст" является установление много-многозначного соответствия глубинно-синтаксических и поверхностно-семантических представлений:

$$(I1) \{ГСП\} \Leftrightarrow \{ПСемП\}$$

В данной работе мы рассматриваем соответствие (I1) с точностью до структур и исключительно в направлении анализа, т.е.

$$(I2) \{ГСС\} \Rightarrow \{ПСемС\}$$

По поводу обратного соответствия  $\{\text{ПСемС}\} \Rightarrow \{\text{ГСС}\}$  в настоящее время нельзя сказать почти ничего определенного сверх того, что оно устанавливается посредством наложения на ПСемС некоторого предложения ПСемС содержательных единиц языка. Автору неизвестен ответ на ключевой вопрос о том, каким способом, отличным от полного перебора, можно выбрать для заданной ПСемС такие семантически содержательные единицы - слова и грамемы, которые с одной стороны, обеспечивали бы полное покрытие смысла ПСемС, а с другой - не выражали бы никакого лишнего смысла.

Установление соответствия (I2) разбивается (не во времени, а чисто логически) на четыре отчетливо выделяемых этапа, каждый из которых обслуживается своим набором правил.

На первом этапе происходит формирование допустимого набора лексических и грамматических значений для каждого из узлов ГСС, или, что то же самое, разрешение лексической и грамматической полисемии и омонимии, т.е. отсеивание ложной полисемии и омонимии.

Значительная часть ложной лексической полисемии и омонимии к моменту перехода  $\text{ГСС} \Rightarrow \text{ПСемС}$  должна оказаться снятой на предшествующих этапах. Так, при интерпретации фразы

(I3) Он сказал мне, что нездоров должно приниматься во внимание только одно - первое - лексическое значение глагола говорить - сказать. Два других значения (О чем вы с ним говорили?, Один мой знакомый бегло говорит по-японски) будут исключены еще на глубинно-морфологическом уровне, поскольку у них нет формы совершенного вида сказать, что отмечается специальной записью в их словарных статьях.

Рассмотрим еще фразу:

(I4) Командный пункт противника находился под наблюдением, в которой нужное (третье) значение слова наблюдение устанавливается с помощью информации о лексических функциях на еще более глубоком уровне представления предложений - глубинно-синтаксическом. Среди прочих лексических функций для лексемы наблюдение 3 определены функции  $\text{Opr}_2$  = находиться (под наблюдением) и  $\text{Adv}_2$  = под (наблюдением), записываемые непосредственно в словарной статье наблюдение 3, а для лексем наблюдение 1 (ср. Наблюдение явлений дифракции позволило ему сформулировать остроумную гипотезу) и наблюдение 2 (ср. Ему принадлежит одно интересное наблюдение) они не определены.

Однако часть ложной лексической полисемии и омонимии и вся граммемная полисемия остается неснятой даже на уровне ГСС. Поясним, что мы имеем в виду.

Как было сказано выше, в узлах ГСС стоят слова глубинно-синтаксического языка с приписанными им семантически содержательными граммемами (грамматическими характеристиками). В общем случае граммема имеет не одно значение, а несколько разных, причем далеко не каждое из них способно комбинироваться со значением лексемы, являющейся именем данного узла.

Граммема несов, например, имеет следующие значения: актуально-длительное (I51), "предстоящее" (I511), узуальное (I5111), потенциальное (I51v), общефактическое, или аористическое (I5v) и другие.

(I51) В соседней комнате кто-то играет "Лунную сонату",

(I511) Завтра он играет в Консерватории "Лунную сонату",

(I5111) Пианист каждый день играет "Лунную сонату",

(I51v) Вы играете на пианино?

(I5v) Кто из участников конкурса еще не играл "Лунную сонату"?

Первое из этих значений не сочетается со значением перформативного глагола в перформативном употреблении. Действительно, фразу

(I61) Клянусь, я не делал этого!

нельзя осмыслить как изображающую процесс клятвы, разворачивающейся во времени. Между тем, в неперформативном употреблении такое осмысление для глагола клясться вполне допустимо, ср.

(I611) Я долго клялся, что ни в чем не виноват, но никто мне не верил.

Возможность реализации актуально-длительного значения грамлемы несов.должна быть исключена именно на этапе перехода от ГСС к ПСемС: ни на каких более ранних этапах этого сделать нельзя, потому что граммема заменяется толкованием (или толкованиями) только в ПСемС.

На втором этапе семантически содержательные элементы ГСС, т.е. лексемы, стоящие в узлах ГСС, и приписанные им грамматические характеристики, заменяются их толкованиями. Весьма существенным при этом является требование, чтобы значения всех содержательных единиц языка, независимо от их

природы, были описаны на одном и том же поверхностно-семантическом языке (см. примечание 2). В противном случае, как уже было сказано, мы будем лишены возможности эксплицировать отношения синонимии между предложениями, построенными из материала разной природы, простым и естественным образом.

Рассмотрим синонимичные предложения

(I7 i) Капитан выстроил батальон буквой П,

(I7 ii) Капитан выстроит батальон в виде < в форме > буквы П.

Значение, которое в (I7 ii) выражено явно - словосочетанием в виде < в форме >, в предложении (I7 i) выражено неявно - творительным падежом существительного. Чтобы продемонстрировать синонимию (I7 i) - (I7 ii) и многих других предложений, в которых падеж существительного семантически содержателен, нужно отказаться от прежнего языка описания грамматических значений и перейти на язык содержательных толкований. В частности, творительный в (I7 i) должен быть истолкован таким образом, чтобы непосредственно в это толкование был включен компонент '...в форме, похожей на вид или форму объекта X'.

На третьем этапе строятся фрагменты ПСемС для пар, троек и т.д. связанных друг с другом значений. Дело в том, что значения отдельных содержательных единиц языка складываются в более крупные блоки далеко не всегда по чисто "аддитивным" законам. Соединяясь друг с другом, значения лексических и грамматических единиц могут подвергаться взаимовлияниям и существенно видоизменяться в различных контекстах. Эти видоизменения должны учитываться особыми правилами взаимодействия.

Рассмотрим предложения

(I8 i) Это ружье не бьет на два километра,

(I8 ii) Неверно, что это ружье бьет на два километра.

Легко видеть, что значения отрицательных предложений (I8 i) и (I8 ii) неидентичны. Второе из них является чисто отрицательным: утверждается, что расстояние, на которое стреляет данное ружье, не равно двум километрам. Оно может быть либо меньше, либо больше двух километров. Между тем, предложение (I8 i) утверждает другую мысль: расстояние, на которое стреляет ружье, строго меньше двух километров. Следовательно, значение отрицательной частицы не, в отличие от значения отрицательного оборота неверно, что, взаи-

действует со значением глагола с количественным дополнением по особому неаддитивному правилу.

На последнем, четвертом этапе в процессе преобразования ГСС  $\Rightarrow$  ПСемС применяются правила-фильтры, с помощью которых устанавливается семантическая правильность ПСемС, построенной на предшествующих трех этапах. Подобно тому, как синтаксическая структура предложения в большинстве естественных языков и в большинстве случаев должна отвечать требованию проективности, ПСемС естественных языков тоже подчиняются ряду универсальных ограничений. Многие такие ограничения указаны в [2]. В частности, в семантически правильной ПСемС 1) пресуппозиционная часть не должна пересекаться с ассертивной; 2) пресуппозиции не должны противоречить ни ассертивной части ПСемС, ни друг другу; 3) содержание модальных рамок тоже не должно противоречить ни остальному семантическому материалу ПСемС, ни друг другу.

Нарушение любого из этих запретов приводит к тому, что соответствующей ПСемС не может быть поставлено в соответствие ни одного правильного в языковом отношении предложения естественного языка. Весьма интересны в этом смысле предложения с эмоционально-усилительными прилагательными, наречиями и частицами, например,

(I9 i) Какой замечательный день!,

(I9 ii) Как хорошо пахнет!

Такие эмоционально-усилительные слова имплицитно утвердительны в том смысле, что их пресуппозицией является утверждение Р, образуемое всем предложением, за вычетом самого эмоционально-усилительного слова. Пресуппозицией (I9 i) является суждение (I9' i), пресуппозицией (I9 ii) - суждение (I9' ii) и т.д.:

(I9' i) День совершенно замечательный,

(I9' ii) Пахнет очень хорошо.

Если придать предложениям (I9 i) и (I9 ii) вопросительную или отрицательную форму, они сразу утратят свою языковую правильность, так как пресуппозиционная часть предложения войдет в противоречие со значением вопросительной интонации (или частицы) или значением не; ср.

(20 i) \*Какой замечательный ли день?,

(20 ii) \*Как хорошо не пахнет!

Подчеркнем, что наличие логического противоречия внутри ассертивной части ПСемС не нарушает ее правильности: такая

ПСемС может быть реализована правильными в языковом отношении предложениями данного естественного языка. Фразы типа

(21 i) Грустная комедия,

(21 ii) Он прожил долгую короткую жизнь

в языковом отношении, конечно, безупречны.

Перечисленные четыре этапа и соответствующие им блоки правил описаны с большей или меньшей степенью подробности в недавно опубликованных работах автора. Кажется, что их достаточно для решения основных задач поверхностно-семантического компонента собственно языковой модели.

Уже сейчас более или менее очевидно, что некоторые другие, преимущественно метаязыковые задачи общей лингвистической теории требуют включения в семантический компонент модели "Смысл  $\Leftrightarrow$  Текст" еще некоторых типов семантической информации и правил, в частности, информации о семантических ассоциациях, или коннотациях данной единицы, а также информации о ее семантических связях с другими содержательными единицами языка. Однако для обеспечения преобразования ГСС  $\Rightarrow$  ПСемС достаточно названных в данной работе четырех типов правил.

Примечания.

I Традиционные грамматики и словари указанным свойством не обладают. Чтобы это утверждение не показалось голословным, сошлемся на следующий хорошо известный факт. В синтаксических описаниях различных языков фигурируют теоретические объекты, которые сейчас принято называть синтаксическими признаками лексем [24], [14]. Одним из относительно простых синтаксических признаков в русском языке является одушевленность - неодушевленность существительных. В некоторых единичных формах (например, в форме винительного множественного) такие существительные по-разному согласуются с прилагательными и причастиями, ср. Вижу белых слонов, но Вижу белые камни. Соответствующее синтаксическое правило согласования должно, следовательно, использовать информацию о том, является ли данное существительное одушевленным или неодушевленным. Очевидно, что перечисление всех неодушевленных и одушевленных существительных непосредственно при правилах согласования, т.е. в грамматике, в высшей степени нерентабельно. Место этой информации - в словаре. Между тем, единственный словарь русского языка, который содержит пометы об одушевленности - неодушевленности существительных - это

грамматический словарь А.А. Зализняка [9]. Ни в одном из существующих толковых словарей, с которыми должны были бы взаимодействовать грамматики, таких сведений нет.

2 Обычно описания значений грамматических и словарных единиц выполняются на разных языках. Значения таких граммем, как ед - мн, описываются с помощью квазисемантических признаков типа 'единичность' (книга), 'цельность', 'неделимость' (виноград), - 'раздельное множество' (книги), 'парность' (боты), 'множественная совокупность' (деньги), 'собираемость' (родители). Между тем, значения словарных единиц не описываются через признаки, а толкуются, например, сокращаться = 'становиться меньше'. Принципиальное различие между описанием значений через признаки и их толкованием состоит в том, что признаковое описание не имеет никакой внутренней структуры и поэтому не может быть осмысленным образом введено в текст вместо определенной единицы, а толкование обладает этим свойством.

3 Некоторые исследователи используют термины "тема" и "данное", "рема" и "новое" в качестве синонимичных. В системе понятий, принятой в модели "Смысл  $\longleftrightarrow$  Текст" [12, 65], эти термины обозначают разные, хотя и связанные друг с другом понятия, поскольку существуют предложения, коммуникативную организацию которых нельзя естественно описать одной парой понятий.

4 Введение составляющих в деревья зависимостей диктуется необходимостью отразить в синтаксических представлениях семантические различия, возникающие из-за того, что определенная единица в составе предложения имеет разные области действия. Предложение Мальчик не болел пять дней имеет два осмысления: (а) 'Мальчик был здоров в течение пяти дней' и (б) 'Мальчик болел менее пяти дней'. В случае (а) областью действия отрицательной частицы не является только глагол, а в случае (б) - глагол вместе с обстоятельством длительности. Эти различия естественней всего удержать посредством синтаксических структур (а')... не  $\leftarrow$  болел  $\rightarrow$  день... (без составляющей) и (б')... не  $\leftarrow$  [болел  $\rightarrow$  день...] (с составляющей). Подробнее см [4].

5 В "падежной грамматике" Ч. Филмора [20], одно время пользовавшейся большим влиянием [28], валентностям предикатных слов приписываются определенные смыслы, или роли - Агэн-



та, Источника, Контрагента, Инструмента и т.п. Похожая система семантически содержательных валентностей была предложена в работе автора [1]. В обзоре [13] корректность подобных семантических языков была поставлена под вопрос. Долгие размышления на эту тему привели нас к выводу, что смысл в семантическом языке, действительно, должен кодироваться единообразно. Если же для его кодирования используются два разных средства - слова и стрелки, оказывается невозможным принципиально ответить на вопрос о том, какие смыслы кодируются словами, а какие - стрелками.

6 Из сделанных выше утверждений отнюдь не следует, что сами видовые значения национально своеобразны. Национальным своеобразием в данном случае обладают не столько значения, сколько средства их выражения. В связи с этим мы бы хотели высказать гипотезу, что арсенал значений, выражаемых глагольными формами русского и французского языков (а также, видимо, других славянских, романских и германских) более или менее совпадает. Различие между русским и французским (или английским) состоит преимущественно в том, что значения, выражаемые по-французски (или по-английски) чрезвычайно богатой и разветвленной системой специализированных форм времени, вида и наклонения, в русском языке специальных форм выражения не имеют. Зато в русском языке в гораздо большей степени развита полисемия глагольных граммем (ср. [5]), когда одна неспециализированная граммема выражает богатейшую гамму видовых, временных и модальных значений. Другое важное различие между русским и французским (или английским) состоит в том, что в русском языке богатая система модально-временных значений, выражаемых глаголом, характеризует преимущественно разговорную речь, а во французском и английском - книжную.

7 Много фактов такого рода собрано в [6], [16]. Большая роль родовых слов во французской лексике объясняется обычно ее абстрактностью [29], а детальность русских видовых номенклатур в области глагола связывается, в частности, с богатством префиксального глагольного словообразования. Явления семантической несоизмеримости разных языков давно стали предметом изучения, а та идея, что язык навязывает человеку выражение определенных мыслей, лежит в основе широко известной гипотезы лингвистической относительности Сэпира-Уорфа (см. прекрасный аналитический обзор [25]).

8 Отметим две конструктивные особенности ПСемС (З'): 1) фрагменты ПСемС, являющиеся образами грамматических значений, содержат абсолютные вершины (узлы, не подчиненные никаким другим узлам). Это формальное свойство "грамматических" фрагментов ПСемС находит вполне естественное объяснение: в русском языке значение любой граммы - это предикат, аргументом которого является значение лексемы и которое само, вследствие независимости грамм разных категорий, не является чьим-либо аргументом. 2) Каждый узел в ПСемС (З') имеет несущий элемент - переменную. При этом условия легко разграничиваются вхождения имени одного и того же референта (одна и та же переменная) и разные референты, имеющие одно и то же имя (разные переменные, ср. XI, X2, X6, X12). Кроме того, становится возможным непринужденное преобразование графа в скобочную алгебраическую запись.

#### Литература

1. Апресян Ю.Д. Лексическая семантика. Синонимические средства языка. М., "Наука", 1974.
2. Апресян Ю.Д. Языковая аномалия и логическое противоречие. - В кн.: *Tekst. Język. Poetyka. Ossolineum, Warszawa, 1978.*
3. Апресян Ю.Д. Принципы семантического описания единиц языка. - В кн.: Уч. зап. Тартуского гос. ун-та. Семантика и представление знаний. Труды по искусственному интеллекту. П. Тарту, 1980.
4. Богуславский И.М. Отрицание в предложениях с обстоятельством в русском языке. - В кн.: *Studia grammatyczne. Ossolineum, Kraków, 1978.*
5. Гловинская М.Я. Видо-временные значения глагольных форм. В кн.: *Русская разговорная речь, т.П (в печати).*
6. Гак В.Г. К проблеме семантической синтагматики. - В кн.: *Проблемы структурной лингвистики 1971. М., "Наука", 1972.*
7. Гладкий А.В. Описание синтаксической структуры предложения с помощью систем синтаксических групп. П. Лингвистическая интерпретация. (В печати).
8. Жолковский А.К., Мельчук И.А. О семантическом синтезе. - *Проблемы кибернетики, вып. 19. 1967.*
9. Зализняк А.А. Грамматический словарь русского языка. Словоизменение. М., "Русский язык", 1977.

10. Иомдин Л.Л. Фрагмент модели русского поверхностного синтаксиса. Оператор согласования для определительных конструкций (прилагательное + существительное). (В печати).
11. Мельчук И.А. К вопросу о грамматическом в языке-посреднике. - Бюл. "Малинный перевод и прикладная лингвистика", вып. 4. М., 1960.
12. Мельчук И.А. Опыт теории лингвистических моделей "Смысл Текст". М., "Наука", 1974.
13. Мельчук И.А. Об одной модели понимания речи (семантическая теория Р. Шенка). I. НТИ, серия 2, № 6, 1974.
14. Мельчук И.А., Перцов Н.В. Модель английского поверхностного синтаксиса. Перечень синтагм. Предварительные публикации Института русского языка АН СССР, вып. 64-66. 1975.
15. Падучева Е.В. О семантике синтаксиса. М., "Наука", 1974.
16. Розенцвейг В.Ю. Языковые контакты. М., "Наука", 1972.
17. Санников В.З. Сочинительные и сравнительные конструкции: их близость, их синтаксическое представление (в печати).
18. Сэпир Э. Язык. Введение в изучение речи. М., Соцэкгиз, 1934.
19. Chomsky, N. Deep structure, surface structure, and semantic interpretation. - In: D.D. Steinberg and L.A. Jakobovits (eds.). "Semantics". London and N.Y., 1971.
20. Fillmore, Ch.J. The case for case. - In: E. Bach and R.T. Harms (eds.). "Universals in linguistic theory". N.Y., Chicago, San Francisco, 1968.
21. Fillmore, Ch.J. Types of lexical information. - In: F. Kiefer (ed.). "Studies in syntax and semantics". Dordrecht - Holland, 1969.
22. Gruber, J.S. Lexical structures in syntax and semantics. North Holland Publishing Company. Amsterdam, N. Y., Oxford, 1976.
23. Jakobson, R. Boas' view of grammatical meaning. "American anthropologist", vol. 61, N<sup>o</sup>. 2, 1959.
24. Katz, J.J. and Postal, P.M. An integrated theory of linguistic descriptions. Cambridge (Mass.), 1964.
25. Kluckhohn, Ch. Notes on some anthropological aspects of communication. "American anthropologist", vol. 63, N<sup>o</sup>. 5, 1961.
26. Lakoff, G. On generative semantics. - In: D.D. Steinberg and L.A. Jakobovits (eds.). "Semantics". London and N. Y., 1971.

27. McCawley J.D. Prelexical syntax. - In: P.Seuren (ed.). "Semantic syntax". London, 1971.
28. Schank, R.S. Conceptual dependency : a theory of natural language understanding. "Cognitive psychology", vol. 3, No.4, 1972.
29. Ullmann, S. Precis de semantique francaise. Berne, 1952.
30. Wierzbicka A. Semantic primitives. Frankfurt, 1972.

ON THE SURFACE SEMANTIC COMPONENT IN THE  
"MEANING-TEXT" MODEL

Ju. Апресян

S u m m a r y

It is argued that the semantic level of sentence representation in the "Meaning-Text" model should be split into two new sublevels - those of surface (national) and deep (universal) semantics. The principal argument in favour of such a division is drawn from the fact that every natural language compels its speakers to express a number of meanings which are irrelevant to the essence of what they mean to convey. This peculiarity is inherent not only in the grammatical system of a natural language, but in its lexical system as well. Surface semantics is responsible for bringing out those language specific meanings and for explicating the relations of synonymy and homonymy between the sentences of the language. Deep semantics is responsible for stripping the surface semantic representations of the fringe of irrelevant meanings and for explicating the relations of synonymy (translatability) between the sentences in all human languages. The surface semantic component of the model maps the set of deep syntactic structures (DSS) onto the set of surface semantic structures (SSemS). The transition DSS  $\Rightarrow$  SSemS is accomplished in four steps: 1) formation of an admissible combination of lexical and grammatical meanings for each node of the given DSS; 2) replacement of each lexeme and each grammeme by its semantic representation; 3) amalgamation of separate lexical and grammatical meanings into larger blocks; 4) inspection of the resulting SSemS for its semantic correctness.

## К СТРАТЕГИИ ПОСТРОЕНИЯ СИСТЕМ КЛАССА "ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ"

В.А. Звегинцев

Лингвистика ныне вовлечена в бурные процессы, вызванные научно-техническим прогрессом, и резко меняет свое лицо. Это, в частности, находит свое выражение в возникновении новых направлений лингвистических исследований, которые во многих случаях не получили еще своего законченного оформления и поэтому не имеют еще достаточно четких границ. Впрочем, общая устремленность этих исследований уже ясно проглядывается, и тем или иным образом связана она с категорией знания - его определения, его становления, передачи, понимания, представления и т.д. Хотя в конечном счете указанные исследования преследуют прикладные, практические цели, они все дальше и дальше отходят от решения возникающих здесь многочисленных задач "в лоб", эмпирически-крохоборческим способом и погружаются все глубже в теорию. Это и логично и вполне оправдано: если исследователь не располагает пониманием и строгим определением тех величин, которыми он оперирует, все его построения неизменно оказываются эфемерными и непродуктивными. Сказанное применительно к той области названных исследований, которая получила условное наименование проблемы искусственного интеллекта, можно проиллюстрировать следующими примерами. Они, кстати говоря, наглядно показывают теоретическую иерархию проводимых здесь исследований и постепенное вычленение конкретных проблем.

На самом высоком теоретическом уровне происходит становление новой науки - науки о мыслительной деятельности человека (Cognitive Science), которую за отсутствием фиксированного термина можно назвать бы когнитологией. Обоснованию необходимости создания такой науки посвящена, например, книга "Репрезентация и понимание" [4], в предисловии к которой ее авторы пишут: "Эта книга содержит исследования в новой области, которую мы именуем наукой о мыслительной деятельности. Данная наука включает элементы психологии, вычислительной математики, лингвистики, философии и педагогики, но она представляет собой большее, чем простое пересечение этих наук. Их интеграция создает новый набор инструментов для ис-

следования широкого круга проблем. В последние годы взаимодействие представителей указанных наук привело к новым и многообещающим результатам в нашем понимании интеллектуальных систем и в развитии науки о мыслительной деятельности. Наша группа научных сотрудников занялась проблемами, которые не представляются разрешенными средствами любой из названных наук в отдельности. Пока слишком рано заниматься предсказанием того, какой курс изберет в будущем это новое сотрудничество, но уже и достигнутое представляется стимулирующим и вдохновляющим" [4; 9-10].

Следующую ступень по теоретической нисходящей, конкретизирующей стратегию столь широко очерченных исследований, представляют недавние высказывания Терри Винограда, который начинал как работник сугубо практического профиля. Свое понимание возникающих здесь проблем он также включает в рамки науки о мыслительной деятельности, но при этом выделяет в ее пределах некоторую область первоочередной теоретической и практической значимости, которую он связывает с так называемым "процедурным" (а фактически динамическим) подходом к лингвистике и особенно к семантике. При этом он поясняет: "В процедурной перспективе "семантика" должна заниматься изучением отношений между лингвистическими объектами и психическими состояниями и процессами, принимающими участие в их создании и осознании" [7]. Исходя из этой установки, Терри Виноград перечисляет положения своей теоретической парадигмы, которые среди прочих включают в качестве основных следующие:

1. При изучении человеческого языка главное внимание следует уделять механизму, лежащему в основе создания и понимания высказываний в лингвистическом и прагматическом контексте.
2. Основные особенности языка отражают мыслительную структуру пользующегося языком человека, включая детальные особенности структур его памяти, обрабатывающие алгоритмы и врожденные вычислительные ограничения.
3. Производство и понимание высказываний имеет место в пределах структуры конкретных мыслительных процессов, включающих как лингвистические, так и нелингвистические элементы. Индивидуальные высказывания могут пониматься только в терминах роли, какую они играют в пределах более широкого контекста. Формализованное понимание языковых процессов является частью более широкой формализации мыслительных процессов.

4. Каждое высказывание строится, чтобы служить комбинации целей, используя набор элементов и возможностей выбора предоставляемых языком" [7; 262] и т.д.

Еще более конкретную программу исследований в области искусственного интеллекта можно найти в книге "Вычислительная семантика". Один из авторов (и редакторов) этого сборника статей - Э. Черняк в своей работе "Вывод и знание" пишет: "Искусственный интеллект имеет дело с созданием компьютерных программ, способных выполнять задачи, которые обычно требуют участия интеллекта... Многие из таких задач включают использование естественного языка, и по этой причине проблема искусственного интеллекта присоединяется к списку дисциплин, предметом которых является изучение природы языка. Вычислительная семантика - наименование, данное нами исследованиям языка, основанным на методах искусственного интеллекта, - использует подход к изучению языка, способный дать ответы на вопросы о том, каким образом используется язык в процессах перевода, при построении вопросно-ответных систем, работающих в режиме языковых текстов и т.д. Этот подход предполагает, что "язык это то, что язык делает" - идея сама по себе не новая для традиционных лингвистических дисциплин, психологии и пр., но вычислительная семантика уникальна в том смысле, что делает эту идею основой для всего изучения языка" [5; 1].

Приведенные высказывания, относящиеся преимущественно к принципам построения систем класса "искусственный интеллект", безоговорочно подчеркивают значение языка, как носителя знаний. Именно это обстоятельство и способствовало, с одной стороны, выдвигению на первый план семантики, а, с другой стороны, и переосмыслению содержания этой дисциплины. Семантика фактически превратилась из области узко лингвистических исследований, занимающихся изучением значений языковых единиц, в дисциплину, выходящую уже за пределы лингвистики и сосредоточивающуюся на изучении того, что именуется человеческими знаниями. Отсюда следуют два фундаментальных вывода. Во-первых, кардинальное изменение статуса лингвистики, которая ныне выступает в качестве науки, имеющей прямое касательство ко всему тому, что имеет дело с человеческими знаниями. Во-вторых, обязательный учет и раскрытие действительной природы генератора знаний - мышления, так же как и достаточно строгое определение его отношений с языком.

Об этом и пойдет в дальнейшем речь, но не в плане общетеоретических рассуждений, а в связи с конкретными и практическими задачами построения систем класса "искусственный интеллект", и, в частности, в аспекте исследовательской стратегии, которой следует придерживаться, по-видимому, на данном этапе решения этой проблемы.

Обо всем этом будет говорить с позиции лингвиста, но сперва представляется необходимым высказывать некоторые общие соображения.

Само наименование исследовательского направления, обозначаемого как "искусственный интеллект" (в США оно даже оформилось в отдельную дисциплину - Artificial Intelligence), безусловно носит метафорический характер. Впрочем, иногда (и довольно часто) это обозначение понимается в буквальном смысле, как проблема, которая должна создать системы, возможно полно имитирующие возможности человеческого интеллекта, и чаще всего именно такое буквальное понимание и является отправной точкой для всякого рода спекулятивных построений. На практике же, как уже отмечалось, дело идет о создании лишь таких систем, которым можно передать некоторые работы, обычно выполняемые человеческим интеллектом. Ставя перед собой такую осторожно формулируемую задачу, мы тем не менее вступаем в совершенно новую научную и общественную эпоху - эпоху интеллектуальных машин. А это событие требует, чтобы к нему отнеслись со всей серьезностью и продуманностью. Оно, бесспорно, обладает и социальным аспектом огромной значимости. Как это в свое время имело место с изобретением письменности.

Находясь в самом начале пути, который ведет нас в новую и во многом еще неизведанную область, не следует задаваться сразу же грандиозными проектами; напротив, надо проявлять больше уважения к запретам, которые налагают на нас наши современные знания и технические возможности. Может быть, надо уделять больше внимания, чем это делалось до сих пор, установлению этих запретов. Но при этом, разумеется, и не упускать из вида конечной цели, без которой все работы класса "искусственный интеллект" неизбежно принимают формы конъюнктурной возни.

Но какова эта конечная цель? Как было уже сказано, практически мы занимаемся не имитированием человеческого интеллекта, а созданием машин, способных выполнять некоторые ра-



боты, исконно рассматриваемые в качестве прерогативы человеческих мыслительных способностей - интеллектуальных машин. С этой точки зрения уже даже и самый непродвинутый компьютер можно назвать интеллектуальной машиной и, если угодно, даже объявить его системой класса "искусственный интеллект". Ведь он осуществляет определенную мыслительную работу, - считает, да еще и получение человека. При таких предпосылках задача, следует думать, заключается в том, чтобы передать машине возможно больше видов работы, которые выполняет человеческий интеллект. Но для того, чтобы иметь основания говорить о такой "разрешающей способности", которая была бы равнозначна человеческой, надо быть человеком. Таким образом, в проблеме искусственного интеллекта, следующего стратегии абсолютной имитации, речь идет фактически не о создании искусственного интеллекта, а о создании искусственного человека.

Совершенно очевидно, что задаваться такой целью мы не можем. Но и стратегия создания последовательности все более и более совершенных интеллектуальных машин - еще не ответ на вопрос, которым мы задаемся. Это - всего лишь путь, которым нам неизбежно приходится идти, но не конечная цель, которую мы можем представить себе пока лишь в приблизительно и в довольно размытом виде. Стоит при этом учесть, что когда человек строил машины, выполняющие "физические" работы, он делал это не для того лишь, чтобы просто имитировать физические возможности своего организма. Видимо, уже с самого начала у него была и иная установка - усилить физические потенции своего организма, и он, как известно, весьма успел в следовании этой установке. Аналогичным образом должно, очевидно, обстоять дело и с интеллектуальными машинами: их назначение - не просто в имитации, а в усилении рабочих возможностей человеческого интеллекта, но - что важно подчеркнуть - не в его огромной и пока непознаваемой целостности, а в отдельных его параметрах.

Названный путь решения проблемы может подать повод и для заключения, что конструированием интеллектуальных машин должны заниматься узкие специалисты по преимуществу технического склада, которые вполне могут обойтись без всяких теоретических осмыслений. Такое впечатление будет неправильным. В действительности все обстоит совершенно противоположным образом, как об этом свидетельствуют и приведенные выше высказывания. Опыт показал, что до тех пор, пока не будут

определены категории, которыми приходится оперировать при создании интеллектуальных машин класса "искусственный интеллект" (примем для удобства изложения этот термин с приведенными выше оговорками), на самом высоком теоретическом уровне, все системы, решаемые исходя из узко конъюнктурных соображений, как правило, бесперспективны. Они могут в какой-то мере удовлетворять сиюминутные потребности, но они не способны создать нужной основы для дальнейшего своего развития и быть использованными в других системах даже того же самого класса. Обычно они и просто неэкономичны.

Поскольку интеллектуальные машины должны исполнять интеллектуальную работу, или, говоря более фигурально, осуществлять мыслительную деятельность, или "мыслить", постольку надо располагать достаточно строгими определениями не только такой величины, как "интеллект", но и многих других, связанных с мыслительными процессами категорий - понимания, знания, мотивации, целеположения, интуиции, научения и т.д. К сожалению, у нас нет такого рода строгих определений всех этих категорий, как и самого понятия интеллекта. "Оказалось, - пишет автор недавно изданной у нас книги Э. Хант, - существует содержательная область знаний, общие принципы которой трудно выделить. Проблема, по-видимому, заключается в определении понятия интеллекта" [3:12]. Но так как такого определения не существует, то психологи, которым и надлежит дать такое определение, используют вместо него "прагматический подход, состоящий в том, что интеллект - это то, что измеряется в интеллектуальных тестах. Так же (пишет Э. Хант) поступлю и я... Принимая этот подход, я откладываю рассмотрение таких нерешенных общих философских вопросов, как "Может ли машина думать?" и "Можно ли говорить, что машина понимает?", до тех пор, пока не будет установлена общая основа фактического знания" [3:12].

Книга Э. Ханта, изданная в США в 1975 г., суммирует опыт уже во многом пройденного этапа в развитии проблемы искусственного интеллекта. Теперь уже стало ясно, что продуктивное развитие этой проблемы просто невозможно без определения названных категорий и этого никак нельзя откладывать. Именно поэтому ныне ученые США и пошли по пути создания упоминающихся выше содружеств из представителей разных наук, способных внести свой вклад в решение данной проблемы. Как пишет Терри Виноград (создатель хорошо известного у нас "робота

Винограда"), интеллектуальным машинам приходится иметь дело со значениями, содержащимися в высказываниях. Однако "не существует какой-либо единой статической структуры, которая могла бы представлять "значение высказывания" и как раз поэтому возникла "наука о деятельности мышления", которая имеет дело с широким кругом мыслительных структур и процессов, включая такие, как восприятие, научение, принятие решений, а также, конечно, язык" [7:263]. Так и выяснилось, что решение казалось бы сугубо практических задач пришлось перенести на высоко теоретический уровень, обращаясь к помощи наук, расположенных на разных классификационных полочках, что между прочим лишний раз говорит не только об условности такого рода классификации, но и разделения на фундаментальные ("чистые") и прикладные науки.

Между тем - и в этом не грех и признаться - многие творцы искусственного интеллекта продолжают следовать в своей деятельности принципу, который Абрахамс Каплан назвал принципом или даже законом инструмента [6]. Этот принцип можно формулировать так: Дайте мальчишке молоток и он будет убежден, что вещи, которые попадутся ему под руку, требуют того, чтобы по ним стучали. Не может вызвать особого удивления то обстоятельство, что ученый формулирует проблемы таким образом, что их разрешение требует как раз той техники, которой он владеет лучше всего. Психологически можно понять такой способ решения проблем, но его - во всяком случае в рассматриваемой нами области - едва ли следует оценить как плодотворный.

Кстати говоря, все изложенное объясняет, почему с такими трудностями встречается установление конечной цели исследований по искусственному интеллекту, о чем говорилось выше. В науке есть такое понятие - пространство логических возможностей. Оно определяется знанием запретов, обусловленных природой изучаемого объекта и выступающих в качестве ограничительных законов. Элементарный пример пространства логических возможностей дает кристаллические многогранники, структура которых лимитируется тридцатью двумя возможными видами их симметрии. В природе найден тридцать один вид и допустимо существование еще одного вида, но два или три вида не может быть - это число уже выходит за пределы пространства логических возможностей. Мы в нашей области исследований сейчас не в состоянии построить такого рода пространство логических

возможностей, так как не обладаем знанием ограничительных законов. А не обладаем мы знанием подобных законов потому, что не знаем природы изучаемых объектов. Здесь представляется уместным сделать несколько самых общих замечаний относительно человеческого интеллекта или мышления, так как несерьезно заниматься построением систем искусственного интеллекта и вместе с тем играть в жмурки с базовым понятием всего этого направления - мышлением. Для любого конкретного исследования нужны хотя бы некоторые ориентиры. Но прежде чем перейти к этим замечаниям, представляется желательным внести ряд уточнений, без чего нередко возникают серьезные недоразумения.

В настоящее время мы располагаем или конструируем автоматические устройства разного порядка. Наибольшего прогресса мы добились в построении таких устройств, которые, используя вычислительную технику и соответствующие программы, оперируют последовательностью операций ("шагов"), цифровыми данными и отношениями, существующими между единицами в той или иной изучаемой системе. Можно сказать, что все подобные устройства находятся на синтагматическом уровне. К ним примыкают и устройства, оперирующие не объектами, а именами объектов. Иными словами, они следуют принципам, принятым в математике, где под формулу  $2 \times 2 = 4$  можно подставить любой объект и никакое из качеств или свойств этих объектов не может оказать влияние на саму формулу. Надо со всей категоричностью подчеркнуть большую важность такого рода устройств для задач определенного класса. Но не даром затем обособилась и возникла проблема искусственного интеллекта, которая, хотя исторически своими истоками и имеет устройства данного порядка, ставит перед собой решение задач уже более сложного порядка, так или иначе соотносимых с многообразными свойствами человеческого интеллекта. Здесь как минимум уже требуется семантическая интерпретация единиц синтагматического уровня и при том в терминах свойств изучаемых объектов. На современном же этапе проблемы искусственного интеллекта учитывается и прагматический уровень.

Во избежание возможных кривотолков необходимо признать, что и системы синтагматического уровня можно при желании соотносить с деятельностью человеческого мышления (об этом уже говорилось выше), но только того вида, который психологи именуют техническим мышлением. Техническое, как и образное (или художественное) мышление, ограничено, однако, достаточ-

но четкими сферами своего применения. Эти виды не идут ни в какое сравнение с высшей формой человеческого мышления - понятийным мышлением, которое прежде всего характеризуется своей универсальностью и которое обычно и имеют в виду, когда говорят о человеческом мышлении вообще (в индивидуальных случаях все эти виды мышления перекрещиваются в разных пропорциях).

Думается, что инерция исторического становления проблемы искусственного интеллекта, восходящая к системам синтагматического класса, все еще оказывает влияние и на понимание основной для проблемы искусственного интеллекта величины - человеческого мышления во всей его полноте. Интеллект, или мышление, стало принятым измерять в лучшем случае мерками логики. И поэтому считается, что если машина способна осуществлять логические операции, она уже есть воплощение человеческого разума. Но характеристики человеческого интеллекта, его мышление никак нельзя сводить к логическим или операционным действиям. Еще раз надо подчеркнуть, что нельзя генерализовать технический аспект мышления. Опираясь на ту же логику, можно, кстати говоря, заключить, что по определению все, что умеет машина, не может относиться к человеческому интеллекту. Человеческий интеллект начинается там, где кончается машина. И, пожалуй, самой большой силой человека является как раз его способность мыслить нелогично и принимать решения поверх всех логических правил мышления. И наконец, самое главное - никакая логика не способна создать нового знания; она направлена на экспликацию знания, но по своей сути тавтологична.

Точно так же едва ли правомерно представлять мышление как функцию одного лишь мозга. Человеческое мышление функционирует не в абстрактной пустоте, повинаясь одним логическим правилам. Его деятельность опирается на всю совокупность данных, которые поставляют ему о "внешнем" мире его сенсорные датчики. Эта деятельность протекает в контексте всего строения человеческого организма, почему и допустимо утверждать, что в процессах мышления принимают участие все клетки человеческого организма, управляясь и координируясь соответствующими центрами. И когда под категорию систем класса "искусственный интеллект" подводят устройства - слепые, глухие, лишенные всех сенсорных датчиков и поэтому находящиеся в полной изоляции от "внешнего" мира, это представляется аб-

сурдным. Как раз преодолению такой изоляции машины и служит естественный язык, конденсирующий в себе все доступные человеку знания об окружающем его мире. И это обязательно должно учитываться при построении систем класса "искусственный интеллект".

Ведь к современным интеллектуальным машинам предъявляют ныне требования умения не только "мыслить", но и общаться с человеком, да еще в режиме естественного языка (что между прочим диктуется и практическими соображениями удобства пользования такими машинами). Иными словами, интеллектуальные машины должны уметь понимать естественный язык и соответственно быть снабжены диалоговыми устройствами, обеспечивающими двухстороннее общение: от человека к машине и от машины к человеку. Конечно, такое общение может осуществляться и помощью "машинного", а не естественного языка. Однако следует помнить: какую бы форму ни принимал такой "машинный" язык, сколько бы он ни был формализован, редуцирован или препарирован любыми способами, он всегда и неизбежно является производным от естественного языка. И от необходимости знания "правил игры" естественного языка не уйти при всех обстоятельствах. А как раз эти правила обычно игнорируются.

Утверждение о связи мышления с языком не пустые или общие слова - оно фактически лежит в основе конструирования машин класса "искусственного интеллекта". А отсюда следует, что все, кто принимают участие в создании таких машин, должны быть лингвистически грамотными специалистами. Ведь интеллектуальные машины оперируют знаниями, а всякое объективное знание лишь постольку знание, поскольку оно выражено на языке. (Более принято говорить в данном случае об информации и соответственно об автоматической обработке информации). Для человека реальность существует совместно с языком. Это вовсе не значит, что язык заступает место реальности, но это значит, что одно всегда вместе с другим. И это значит, что знания, фиксируемые языком, в языковой форме поступают и в распоряжение машины.

Для многих исследователей, работающих в интересующей нас области и осознающих важность лингвистических данных, язык и его единицы в лучшем случае интересны в первую очередь как носители "значений". Отсюда и стремление сформулировать семантические представления (или репрезентации) высказываний, назначение которых состоит не только в том, чтобы избежать

многозначности высказываний, но и по возможности освободиться от ограничений, налагаемых языком, от всех тех многочисленных особенностей, которые свойственны естественному языку и которые кажутся лишними и явно мешают машинной обработке информации. Такой путь предполагает сознательную редукцию целей создаваемой системы и резкое уменьшение ее эффективности. Для современных специалистов в области искусственного интеллекта такая постановка проблемы представляется уже неудовлетворительной. Для них "значение есть многомерная величина, формализуемая только в терминах целостностного комплекса целей и знаний, соотносимых как с продуцируемым их, так и с воспринимающим и понимающим их" [7:262]. Эта довольно трудно воспринимаемая формула требует своего раскрытия и корректировки.

"Создание" языковых значений есть акт познания. Поскольку это так, постольку такого рода "создание", осуществляемое мышлением, опирается на данные, поставляемые человеку его сенсорными датчиками. И если бы у человека они были бы иными, чем есть, то и значения слов естественного языка были бы также иными. Но это еще не все. Надо знать и непременно учитывать, что за каждой единицей языка стоит вся система языка, и она постольку является действительно языковой единицей, постольку входит в эту систему. Вне языковой системы не существует его единиц, так во всяком случае обстоит дело в естественном языке, от которого, как уже было сказано выше, хочешь или не хочешь, а неизбежно приходится танцевать.

Но и это еще не все. Язык - не арифметическая совокупность слов. Язык фактически начинается с предложений. Вот почему ныне конструкторы искусственного интеллекта предпочитают иметь дело как раз с предложениями или даже с текстами.

Посредством предложений мы не только передаем нашим собеседникам некоторую совокупность "информации". Предложение выполняет чрезвычайно многообразные функции и обладает весьма сложной природой, что, как правило, также игнорируется при построении систем, имеющих, в частности, диалоговое устройство. Предложение выражает суждение, служит определенным целям в речевом акте, обнаруживает отношение к сообщаемому (т.е. имеет те или иные модальности) и к тому, кому адресуется предложение, учитывает условия, в которых осуществляется речевой акт, и т.д. Более того, строя предложение (а точнее высказывание), мы определенным образом структурируем

окружающую нас действительность, потому что построение предложения предполагает вычленение из континуума действительности некоторого ее фрагмента-ситуации, - которая и становится содержанием предложения, его "значением". С полным основанием можно утверждать, что во многом с помощью предложения человек ориентируется в окружающем мире.

А еще дальше лежит еще совсем не опознанная область - связный текст, который и является подлинной реальностью языковой деятельности человека. Ныне эта область осваивается лингвистикой текста, которая делает только первые шаги. Оказавшись на еще совершенно неизведанной территории, лингвистика текста с самого начала столкнулась с большими трудностями, что и привело к чрезвычайной разноголосице мнений и точек зрения. Практическую значимость исследований в данной области отлично характеризует тот факт, что именно при выходе за пределы предложения споткнулся машинный перевод.

И, наконец, пожалуй самое существенное. В свое время А.М. Пешковский писал: "... Точность и легкость понимания растут по мере уменьшения словесного состава фразы и увеличения ее бессловесной подпочвы. Чем меньше слов, тем меньше недоразумений" [2:297]. Традиционная лингвистика является наукой о "словесном составе фразы", она строит "грамматики произнесения". Но поскольку речь идет о деятельности общения, которая невозможна без понимания, поскольку "грамматика произнесения" должна быть пополнена "грамматикой умолчания". Несказанное, не получившее словесного выражения нередко бывает более содержательным и более красноречивым, чем "словесный состав". И совершенно очевидно, что "умолчания", направленные на большую легкость понимания, подчиняются строгим закономерностям, которые еще совершенно не изучены нами. Пожалуй, единственное, что мы знаем о "грамматике умолчания" сводится к тому, что нарушение правил этой грамматики приводит к нарушению процесса общения, делает этот процесс бессмысленным [1].

Это, разумеется, далеко не все сложности, которые связаны с языком. Но и то, о чем по необходимости упоминается лишь в самом общем виде, достаточно наглядно показывает, с чем неизбежно придется столкнуться исследователю, если он задается задачей построить полноценную интеллектуальную машину, способную эффективно выполнять не слишком ограниченный и узкий круг работ. Следует еще раз подчеркнуть, что совре-



менный этап исследований в области искусственного интеллекта исходит именно из этой позиции.

Все изложенное, очевидно, дает основание для следующих заключений. Построение систем класса "искусственный интеллект" имеет преимущественно практическую направленность и имеет целью передать машине выполнение некоторых работ, требующих участия человеческой мыслительной деятельности и связанных со знанием - как бы это последнее не определялось. В силу этого более правомерно говорить не о проблеме искусственного интеллекта, а о стратегии построения интеллектуальных машин. Эта стратегия обуславливается не только техническими ограничениями, но в еще большей степени пониманием и достаточно строгим определением как основных параметров мыслительной деятельности, так и познанием природы знания, имеющих в равной мере своим основанием язык. Для объединения столь разнородных компонентов необходимо создание (следует думать, в рамках науки о мыслительной деятельности) специальной теории - достаточно сильной, чтобы возникло целостное образование, а не механическая совокупность разнородных данных. Значение и ответственность этой теории находится в прямой зависимости от того обстоятельства, что колоссальное усиление с помощью машин и соответствующих программ мощи отдельных аспектов мыслительной деятельности человека может привести к возникновению ряда совершенно новых проблем и возможно - с чем также надо считаться - к непредвиденным результатам.

#### Литература

1. Звегинцев В.А. К вопросу о природе языка. - "Вопросы философии", 1979, № II.
2. Пешковский А.М. Объективная и нормативная точки зрения на язык. Цит. по книге В.А. Звегинцева "История языкознания XIX и XX веков", т. 2, М., 1965.
3. Хант Э. Искусственный интеллект. М., 1978.
4. Bobrow, D. and Collins, A. (eds.). Representation and Understanding. N.Y. - San Franc. - Lond., 1975.
5. Charniak, E. and Wilks, Y. (eds.) Computational Semantics. Amsterdam-New York - Oxford, 1978.

6. Kaplan, A., The Conduct of Inquiry. San Francisco, 1964.
7. Winograd, T., Towards a Procedural Understanding of Semantics. "Revue internationale de Philosophie", Brussels, 1976.

SOME NOTES ON THE STRATEGY OF BUILDING SYSTEMS  
OF THE "ARTIFICIAL INTELLIGENCE" CLASS

V. Zvegintsev

S u m m a r y

Linguistics is going through a rapid process caused by the scientific-technological revolution, and is speedily changing its face. This process finds its expression in the appearance of new trends that have not yet found their final make-up and that lack precisely delimited boundaries. All these trends are in some way or other connected with the category of knowledge. Although the present investigations are of a practical and applied character they should be based more on theoretical studies as the investigator should understand and define the concepts being operated with precisely. Artificial intelligence investigations may serve to corroborate this assertion.

On the highest theoretical level the groundwork of a new branch of science - a science of human intellectual activities (cognitive science) - is laid down. The next level is represented by theoretical and practical tasks singled out from the general artificial intelligence problem domain by T. Winograd. An even more concrete program of investigations is proposed by E. Charniak and Y. Wilks, editors and co-authors of the book "Computational Semantics."

The author of the paper claims that the development of the systems of the artificial intelligence class is primarily concerned with practical problems and that its aim is to transfer to machines the fulfillment of some tasks demanding human intellectual abilities and connected with

knowledge. The author underlines that it would be more proper to speak not of the problem of artificial intelligence but rather of the strategy of building intellectual systems. This strategy is conditioned not only by technical restrictions but to a greater degree by understanding and determining quite strictly both the main parameters of intellectual activities and the cognition of the essence of knowledge. To unite such different components a specific theory is necessary that should be strong enough to create an integral complex, not a mere collection of individual components.

## ПОСТРОЕНИЕ РЕДУЦИРОВАННОГО ЯЗЫКА ДЛЯ АВТОМАТИЧЕСКИХ ДИАЛоговых СИСТЕМ

О.В. Звегинцева

В настоящей работе будут иметься в виду диалоговые системы, обеспечивающие общение между человеком и машиной, работающие в режиме естественного языка и ориентированные на определенную предметно-понятийную область (проблемную среду). Последнее обстоятельство, кстати говоря, во многом разрешает проблему омонимии, которая является камнем преткновения при разработке программ общения между человеком и машиной. Даже и профессиональные языки, которые можно рассматривать как первичную форму редукции естественного языка, в первую очередь характеризуются однозначностью употребления своей терминологии.

Уже и сама ориентация на определенную проблемную среду по сути дела предполагает осуществление определенной редукции естественного языка, поскольку проблемная среда представляет всего лишь некоторый фрагмент действительности, в пределах которого происходит данная конкретная деятельность человека, совершаемая им совместно с машиной. Границы проблемной среды могут варьироваться в чрезвычайно широких пределах и соответственно изменяется степень редукции языка.

Но в практике построения диалоговых систем следует учитывать обусловленность редукции и иного порядка - исходящую из рабочих возможностей машины (компьютера) и используемой программы. Ведь все и даже самые совершенные системы класса "искусственный интеллект", работающие в режиме естественного языка, используют пока ограниченный - и при том сильно ограниченный - естественный язык. И как раз в постепенном снятии ограничений на естественный язык следует видеть один из главных способов совершенствования диалоговых систем.

Учитывая данные предпосылки, важно разработать общие принципы и методику осуществления редукции естественного языка. Совершенно при этом очевидно, что здесь нельзя следовать принципам, которым следовали создатели упрощенных языков типа Basic English. Такого рода упрощенные языки неприемлемы уже по той причине, что они рассчитаны на универсальное потребление и не учитывают специфики проблемной среды,

на обслуживание которой и ориентируется в той или иной степени построение диалоговых систем. Именно проблемная среда и ее специфика и должна выступать в качестве определяющего фактора при проведении редукции языка. Иными словами процесс редукции должен идти не от естественного языка как такового, а от того конкретного корпуса действительного употребления языка, который используется в данной конкретной среде.

При таком подходе относительно просто установить лексические ограничения. При необходимости, если объем используемой лексики все же превосходит рабочие возможности машины и программы, здесь допустимо проведение дополнительных ограничений, базирующихся на частотности употребления отдельных лексических единиц. Более сложно обстоит дело с синтаксическими конструкциями. Их набор, конечно, также извлекается из данного конкретного корпуса, используемого в данной проблемной среде фрагмента естественного языка. Но дело не ограничивается лишь определением того, что для данного языка является необходимым, и что излишним и неупотребительным. Это, разумеется, также нужно знать для того, чтобы дать достаточно исчерпывающее описание данного ограниченного языка. Однако не менее существенным является и установление того, что в таком языке опускается и восполняется "знаниями того мира", в котором происходят описываемые на используемом языке события. Следует учесть, что без этого оказавшегося за пределами текста "знания" невозможно правильное его понимание.

Ниже приводится пример анализа синтаксических особенностей текста в описанных выше целях. В качестве анализируемого текста берутся "Правила и фразеология радиообмена между экипажами воздушных судов и диспетчерами служб движения гражданской авиацией". При этом принимается, что эти "Правила" фактически используют редуцированный язык, который и подлежит процедуре экспликации, поскольку остаются неизвестными его конкретные ограничения и, следовательно, действительный набор синтаксических конструкций, лежащих в основе построения программ синтеза и анализа для машинного их использования.

Нормализуемой единицей в данном тексте является реплика. Репликой является одно (или более) предложение, приписываемое "Правилами" одному говорящему. Все реплики классифицируются по принципу схожести синтаксической структуры. Язык радиообмена задается в виде таблиц, семантика этого языка

определяется посредством перевода реплик на естественный язык. Таким образом, для каждой реплики мы имеем описание на 2-х языках - редуцированном языке радиообмена и естественном языке. Соответственно выясняются и контекстные особенности редуцированного текста и свойственные ему опущения. Контекстные опущения и общие синтаксические особенности определяются, следовательно, посредством сравнения наличия/отсутствия различных категорий, встречающихся в редуцированном тексте, с Академической Грамматикой русского языка, являющийся наиболее полным каталогом.

В данной статье рассматриваются наиболее общие и типичные конструкции реплик. Для формализованного представления этих конструкций вводятся некоторые обозначения:

V	- глагол
N	- существительное
Adv	- наречие
Prep[...]	- предлог
Adj	- прилагательное
Pron	- местоимение
Sing	- единственное число
Pl	- множественное число
temp	- категория времени
loc	- обстоятельства места
inf	- инфинитив
*	- факультативность

Нижние индексы при N обозначают падежи

Перечень редуцированных конструкций:

I. V I. л., Sing, наст. после которого могут следовать

$$1) N_{вл.}(\underbrace{Prep \{на\} N_{л.о.}}_{loc})^*$$

"Разрешаю пересечение";

"Разрешаю запуск на РД..."

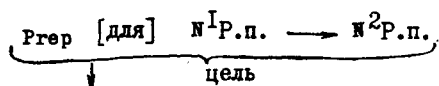
2) Adj сравн. степ. (Adj субстантивированное)

"Разрешаю правее"

"Разрешаю исполнительный".

При V I. л., Sing наст. вр. опускается прямое дополнение, выраженное существительным.

# 1) Вместо группы



Прер [на] N<sup>3</sup>

ставится группа Прер [на] N<sup>3</sup>, т.е. убирается часть со значением цели.

"Разрешая буксировку для запуска двигателей на РД № .."  
- "Разрешая запуск на РД..."

2) Вместо Adj сравн. степ., → N Р.п. ставится Adj сравн. степ., т.е. убирается существительное, ясное из контекста (при ситуации "в районе аэродрома взлета", например, таким существительным является "трасса").

Вместо Adj ← N ставится Adj субст., т.е. тоже убирается существительное.

"Разрешая обход правее трассы" -

"Разрешая правее".

П. Это тот случай, когда конструкция реплики охватывает более чем одно предложение.

Предл. 1: V прош., Sing; Предл. 2: N им.п.

Как правило, эта конструкция имеет одно лексическое наполнение: "Понял. Курс..." Однако Предл. 1 и Предл. 2 могут встречаться и по отдельности.

При таком типе редукции опускается группа цели (обычно: Прер [для] N Р.п.) и глагол (I л., Sing, наст./прош.), прямым дополнением которого является в полной фразе существительное, остающееся при редукции.

"Для обхода стаи птиц взял курс..." -

"Понял. Курс..."

Ш. В этом типе - предложения, содержащие указание на место, на котором должен оставаться, или к которому должен двигаться адресат. Обычно сначала указывается действие (напр., "перелет"), а затем место. Действие может быть опущено. Если же положение не изменяется, то употребляется временное наречие

1) N им. п.

2) Adj<sub>temp</sub>

3) Ø

Прер [на] ... (loc)

"Перелет на ... (север)"

"Пока на ... (эшелон)"

"На... (контрольный пункт коридора)"

1) Вместо группы слов берется одно, выражающее идею этой группы (напр.: "птицы смещаются на..." - "перелет").

2) Вместо глагола (повелит., Pl) + Adj<sub>temp</sub> берется Adj<sub>temp</sub>, но не то, которое стояло с глаголом, а другое того же значения ("временно" - "пока").

3) Опускается указание действия, ясного и однозначно восстановимого в данной ситуации.

IV. V I л., Sing, наст. ... (Н В.п.)

Эта конструкция состоит из глагола I л., ед. ч., наст. времени, у которого может быть факультативно прямое дополнение.

"Уточняю пролет"

"Пересекаю ... (эшелон)"

Обычно Н В.п., зависящее от V I л., Sing, наст., опускается. Опускается также уточнение (напр., временное - "Сейчас")

"Сейчас пересекаю эшелон..."

"Пересекаю... (эшелон)"

У. Adj субст. ... (Prep [на] )\*

Конструкция состоит из субстантивированного прилагательного, после которого, при указании данных местонахождения, стоит предлог "на"

"Встречный на..."

Опускается существительное, при котором было прилагательное, ставшее субстантивированным в редуцированной конструкции.

"Ваш встречный борт на... (эшелон)" -

"Встречный на ... (эшелон)"

Опущено также Prep "Ваш".

VI. ... (XI им.п., N им.п.)

Такие конструкции встречаются при сообщении данных, когда в скобках ставятся имена, к которым данные относятся

"... (эшелон, высота)"

В редуцированной конструкции при существительных имеются данные, уточняющие выраженные существительными денотаты. Опускается уточнение действия, ясное из ситуации, и, если при глаголе есть уточнения времени, места, ясные из контекста.

"Занял эшелон..., выход на трассу через (коридор)"

"... (эшелон) на ... (контрольный пункт коридора)"



УП. (и В.п.)<sup>\*</sup> ←  $v_{inf}$

Обычно это либо приказ ждать дальнейшей информации ("ждать"), тогда нет прямого дополнения, либо требование доклада ("Курс доложить")

В редуцированной конструкции остается инфинитив, имеющий значение повеления, и при нем может быть дополнение, выраженное существительным. Опускается стандартное действие, производимое с денотатом существительного и уточнение образа действия.

"Выполняйте набор эшелона... по установленной схеме, пересечение эшелона... доложить" -

"Набор... (эшелон), пересечение... (эшелон) доложить".

УШ. ... и В.п.

v повелит., Р1

... Прер [с] и Т.п.

Это приказ к действию, при этом указывается либо объект действия, либо партнер.

"Занимайте... (эшелон)"

"Работайте с Центром".

Стандартное действие, выраженное глаголом (повелит., Р1) + пациенс, выраженный существительным (В.п.) + уточняющее пациенс слово заменяются на глагол, выражающий идею группы + пациенс + уточняющее слово (существительное)

"Выполняйте набор эшелона... с МПУ..." -

"Занимайте... (эшелон) с курсом..."

В результате проделанного анализа оказывается возможным установить в самом общем виде особенности рассматриваемого редуцированного языка (ради экономичности изложения приходится отказаться от более полного их перечисления).

Эти общие особенности заключаются в следующем:

- 1) глаголы не употребляются в форме сослагательного наклонения;
- 2) у глаголов отсутствует форма 2 лица, единственного числа и 1 лица, мн.ч.;
- 3) прилагательные не употребляются в превосходной степени;
- 4) не употребляются причастия и деепричастия;
- 5) нет словосочетаний с субъектным инфинитивом (типа "решил остаться", "умеет жить");
- 6) нет сложноподчиненных предложений;
- 7) нет предложений, где подлежащее было бы выражено ин-

финитивом или словосочетанием (имя + имя);

8) не употребляется составное именное сказуемое;

9) нет обособленных второстепенных членов предложения и цепочек определений.

Для языка радиообмена характерна повышенная употребляемость повелительного наклонения. С поверхности реплик убраны выражения межфразовых связей, хотя иногда наблюдается наличие в одной реплике нескольких предложений.

Как было сказано выше, полное и правильное понимание высказываний (реплик) достигается за счет привлечения "знаний о мире" данной проблемной среды. Это делается с помощью введения в память машины ситуативного каркаса проблемной среды, при поступлении которого в качестве основных ориентиров выступают опущения, выявляемые при анализе структуры реплик. Методика и способы построения такого рода ситуативного каркаса проблемной среды - особая проблема.

# CREATING A REDUCED LANGUAGE FOR AUTOMATIC DIALOG SYSTEMS

O. Zvegintseva

## S u m m a r y

A natural language based problem-oriented man-machine dialog system is considered. The use of language in a fixed problem domain already presupposes substantial reduction of natural language; at the same time there exists also a second possibility of reduction arising from the working limitations of the computers and programs used. Therefore it is considered necessary to work out the principles and methodology of substantial reduction of natural language. In so doing, the author says that the process of reduction should proceed not from natural language as such but from the concrete use of language in the given problem domain. As an example she considers "Rules and phraseology of the exchange of radio messages between air craft crews and flight controllers in civil aviation". The basic unit of such a type of text is the replay. The author examines only the most common and typical of them. She states that complete and correct understanding of a replay can be achieved by involving "knowledge of the world" in the given problem domain. This is materialized by introducing into a machine's memory the situational frame of the problem domain.

# К ПОСТРОЕНИЮ ФОРМАЛЬНОЙ МОДЕЛИ СОВРЕМЕННОГО РУССКОГО ЯЗЫКА. ОПЕРАТОР СИНТАКСИЧЕСКОГО СОГЛАСОВАНИЯ ДЛЯ ОПРЕДЕЛИТЕЛЬНЫХ КОНСТРУКЦИЙ

Л.Л. Иомдин

В различных системах искусственного интеллекта, ориентированных на использование естественного языка (взаимодействие человека и ЭВМ на естественном языке, информационно-поисковые и управляющие системы, автоматическое индексирование и реферирование, машинный перевод и др.), одной из самых принципиальных проблем является создание надежных действующих моделей естественного языка и отдельных фрагментов таких моделей.

В настоящей работе предлагается фрагмент формальной модели современного русского языка - оператор синтаксического согласования  $СОГЛ_A(S)$ , посредством которого задается описание одного из основных типов русского согласования - согласование адъектива<sup>I</sup> (прилагательного и причастия) с существительным, имеющего место в определительных конструкциях типа высокая сосна, два старых пня, поваленные ольха и береза, белый и подосиновый грибок, а также в ряде других конструкций. Оператор  $СОГЛ_A(S)$ , построенный на теоретических принципах модели "Смысл  $\longleftrightarrow$  Текст" [9], представляет собой набор правил, допускающих непосредственную машинную реализацию, и может функционировать в качестве автономного механизма в системах автоматического синтаксического анализа и синтеза русских текстов.

На первый взгляд может показаться, что согласование указанного типа - это достаточно локальное явление, которое затрагивает лишь синтаксический уровень языка и может быть описано исключительно в синтаксических терминах. На самом деле, однако, это согласование носит весьма нетривиальный характер и имеет прямое отношение не только к синтаксису, но и к семантике языка; в частности, обнаруживается, что грам-

<sup>I</sup> В дальнейшем, для простоты, мы будем пользоваться термином "прилагательное" для обозначения как собственно прилагательных, так и причастий - в тех случаях, когда это не ведет к недоразумению.

матические характеристики согласуемых слов в некоторых случаях несут значительную смысловую нагрузку и должны по этой причине учитываться особым образом. Именно такой подход принимается в настоящей статье.

По методике изложения и способам представления языковых выражений данная статья тесно примыкает к статьям [7] и [8] и может считаться их продолжением.

## § I. Предварительные замечания

В процессе поверхностно-синтаксического анализа и синтеза русского текста часто оказывается необходимым учитывать то или иное соответствие, или с о г л а с о в а н и е, морфологических (а иногда и словарных) характеристик лексем - членов каких-либо синтаксических конструкций. Разные типы такого согласования в содержательных терминах описаны в грамматиках и хорошо известны; тем не менее нам представляется полезным, прежде чем приступить к формулировке правил оператора, напомнить основные типы согласования в современном русском языке.

### I. Согласование личного глагола с существительным

Глагол в неперошедшем времени обычно согласуется по числу с неместоименным существительным (мальчик бежит, но не \*мальчик бегут); по лицу и числу - с местоименным существительным (ты бежишь, но не \*ты бегу < бежите >); глагол в прошедшем времени согласуется с существительным по роду (род - словарная характеристика существительного!) и числу (мальчик бежал, но не \*мальчик бежала < бежали >).

### 2. Согласование прилагательного с существительным

Здесь удобно различать три случая:

а) полное согласование - прилагательное согласуется с существительным по роду, числу, падежу и одушевленности (высокая сосна, но не \*высокий < высокие, высокой > сосна; [снял] белого ферзя, но не \*... белый ферзя);

б) частичное согласование - прилагательное согласуется с существительным по роду и числу (объявить собрание закрытым, но не \* объявить собрание закрытой < закрытыми >; ночь нежна, но не \*ночь нежен < нежны >);

в) частичное согласование - прилагательное согласуется с

существительным по роду и одушевленности (мальчики, каждого [из которых я знал], но не \*мальчики, каждую < каждый > из которых я знал; [послали лучшего < лучших > из студентов, но не \*... лучшую < лучшие из студентов).

### 3. Согласование числительного с существительным

а) по падежу (пяти < двум > девочкам, но не \*пять < две > девочкам);

б) по роду (две девочки, но не \*два девочки);

в) по одушевленности (двух девочек, но не \*две девочек);

г) по роду, числу, падежу и одушевленности (оштрафовали одного пассажира, но не \*одну < одних, один > пассажира).

### 4. Согласование двух существительных

а) по падежу ([купил] участок и дом < участок и дома >);

б) по падежу и числу (студента-медика).

Как правило, один и тот же тип согласования характеризует несколько разных синтаксических конструкций. Например,

- согласование глагола  $\gamma$  с существительным  $X$  обнаруживается в предикативных конструкциях (подлежащее + сказуемое), ср. предикативное

Маша [ $X$  = (жен), ед] была [ $\gamma$  = жен. ед.] здесь<sup>2</sup>, а также в некоторых комплетивных конструкциях (связка + имен-

I комплетивное

ная часть сказуемого), ср. Это была [ $\gamma$ ] Маша [ $X$ ];

- полное согласование прилагательного  $\gamma$  с существительным  $X$  (тип 2а) - в адъективных конструкциях, описываемых с помощью нескольких типов синтаксических отношений в т.ч. определительного (хорошая [ $\gamma$ ] книга [ $X$ ]), описательно-определятельного (тяжело раненного [ $\gamma$ ] солдата [ $X$ ] отправили в тыл), аппроксимативно-порядкового (страницы  $X$  с десятой [ $\gamma$ ]), а также в некоторых других, более сложных конструкциях атрибутивное

(о предмете [ $X$ ] как подч-союз → таковом [ $\gamma$ ], говорить → о → задаче [ $X$ ] как → о → решенной [ $\gamma$ ], мужчины [ $X$ ] высокого роста → и → низкорослые [ $\gamma$ ]);

<sup>2</sup> Здесь и далее стрелки обозначают направление синтаксической зависимости, а над стрелками указываются название соответствующих поверхностно-синтаксических отношений (см. о них [7] - [9]).

- частичное согласование прилагательного Y с существительным X (тип 2б) - в некоторых комплетивных (он [X = (муж),

I компл  
ед] стал старым [Y = муж, ед]), субъектно-копреди-  
кативных (он [X] вернулся суб-копр старым [Y]) и объект-  
но-копредикативных конструкциях (его [X] помнили об-копр  
старым [Y])<sup>3</sup>;

- согласование числительного Y с существительным X (тип 3) - в количественных (десяти колич студентам) и аппроксими-  
мативно-количественных конструкциях (студентам аппрокс-колич  
десяти).

Как известно, согласование лексем - членов некоторой конструкции зависит от целого ряда контекстных условий, а также от словарных особенностей самих этих лексем и формулируется достаточно сложным образом. Иной раз соответствие морфологических и синтаксических характеристик бывает настолько запутанным, что от согласования в узком смысле слова - совпадения значений одноименных морфологических категорий

предик  
- не остается и следа: ср. он (3 л. ед) придет (3 л. ед), но  
предик  
он (3 л. ед) и ты придете (2 л. мн); высокая (жен.ед. им)  
опред  
сосна ((жен). ед. им), но две высокие (мн. им) сосны (ед.  
род)<sup>4</sup>.

<sup>3</sup> Подчеркнем, что в похожих конструкциях с существительными (он стал вернулся стариком), на наш взгляд, нет синтаксического согласования - ни по роду, ни по числу. Грамматическая неправильность фраз типа Он стал старой или Он стал стариками, в отличие от Он стал старой <старыми>, обусловлены нарушением семантического, а не синтаксического согласования; с некоторой натяжкой можно было бы считать, что эти фразы выражают абсурдные смыслы типа некий мужчина стал старой женщиной или некий мужчина стал старыми мужчинами, хотя такие "смыслы" следовало бы выразить более эксплицитным образом (о семантическом согласовании и о степени эксплицитности выражения неожиданных смыслов см. [1]) - ср. абсолютно нормальные и осмысленные предложения типа Он стал старой бабой или Этот океанский лайнер стал жалкими олодомками металла. Согласование двух существительных по роду/полу, вероятно, вообще не встречается в русском языке.

<sup>4</sup> Отмечая такую терминологическую некорректность, мы все же предпочитаем использовать привычный и достаточно удобный термин "согласование" для описания всех случаев соответствия морфологических и синтаксических характеристик членов конструкций определенных типов. Возникающие здесь теоретические проблемы (в первую очередь, серьезную проблему "согласование или управление" - см. о ней [2], [6]), мы сознательно оставляем в стороне.

Учет таких условий непосредственно в синтаксических правилах - синтагмах привел бы к неоправданному увеличению их числа и одновременно к потере наглядности. Представляется поэтому целесообразным иметь в системе поверхностного синтаксиса особые механизмы, которые позволяли бы учитывать необходимую информацию о соответствии характеристик членов тех или иных синтаксических конструкций независимо от синтагм - операторов согласования (см. [7], [9]).

В системе русского поверхностного синтаксиса используется несколько таких операторов. В частности, соответствие характеристик глагола и существительного обеспечивается оператором  $COPL\ y(s)$ , характеристик прилагательного и существительного (согласование типа 2а) - оператором  $COPL_A(s)$  и т.д. При анализе предложения эти операторы проверяют, выполняется ли в глубинно-морфологическом представлении ЭГМП этого предложения требуемое соответствие характеристик лексем - кандидатов в члены какой-либо конструкции; при синтезе же посредством операторов согласования лексемам - членам данной конструкции - приписываются недостающие морфологические характеристики.

Отметим две следующие важные особенности операторов согласования.

Во-первых, все операторы согласования формулируются отдельно от синтаксической конструкции, содержащей лексему, морфологические характеристики или синтактики которых обрабатываются данным оператором: один и тот же оператор можно применять для описания разных конструкций, если между характеристиками их членов имеется однотипное соответствие. Операторы согласования, тем самым, определяются просто для пар словоформ лексем разных синтаксических классов (обычно частей речи).

Во-вторых, операторы согласования являются направленными: предполагается, например, что прилагательное согласуется с существительным в определенной конструкции, а не существительное с прилагательным; что глагол согласуется с существительным в предикативной конструкции, а не наоборот. На практике это означает, что при анализе некоторые компоненты морфологической характеристики согласуемого слова  $\bar{Y}$ , фигурирующие в ГМП, опускаются в поверхностно-синтаксическом представлении (ПСП): считается, что они заданы характеристиками слова  $X$ , с которым  $\bar{Y}$  согласуется; при синтезе же, наоборот,



отсутствующие в ПСП компоненты морфологической характеристики  $\bar{Y}$  -а восстанавливаются по  $X$ -у в ГМП.

Отметим, что направление согласования, вообще говоря, не имеет отношения к направлению синтаксической зависимости (ср., например, конструкции "глагол + существительное", где глагол, будучи синтаксической вершиной, является согласуемым словом); более того, между членами согласовательной пары прямой синтаксической связи может вообще не быть (см. выше примеры на согласование типов 2б и 2в).

## § 2. Оператор согласования $\text{СОГЛ}_A(s)$

Ниже предлагается описание оператора  $\text{СОГЛ}_A(s)$  ( $\bar{Y}$ ,  $X$ ), задающего соответствие морфологических и синтаксических характеристик прилагательного  $\bar{Y}$  и существительного  $X$  в русских адъективных конструкциях (согласование типа 2а).

Сначала - несколько неформальных замечаний.

Мы уже говорили, что при определенных условиях характеристики согласуемых слов могут соотноситься более сложным образом, чем в "чистом" случае, когда согласование  $\bar{Y}$  -а с  $X$ -ом сводится к совпадению значений одноименных грамматических категорий.

Отклонения от стандартного случая вызываются в первую очередь следующими причинами.

1. Несовпадение объема значений одноименных грамматических категорий существительного и прилагательного. Сюда относится явление падежного "рассогласования"  $X$ -а и  $\bar{Y}$ -а в тех случаях, когда  $X$  - в партитивном или местном падеже; ср. выпил крепкого [ $\bar{Y}$  = род] чаю [ $X$  = парт]; в основном [ $\bar{Y}$  = предл] бору [ $X$  = местн].

2. Синтаксические особенности существительного  $X$  позволяют ему иметь определение  $\bar{Y}$  в другом роде или числе; ср. наша [ $\bar{Y}_1$  = жен] знакомая [ $\bar{Y}_2$  = жен] врач [ $X$  = (муж)], эти [мн] миллион [ед] рублей.

3. Существительное  $X$  имеет при себе в качестве зависимого "малое" числительное полтора, два, оба, три, четыре в именительном или винительном падеже; ср. два [им] добрых [ $\bar{Y}$  = мн. род] друга [ед, род], [видал лишь] четыре [вин] голые [мн. вин] стены [ед. род].

4. Существительное  $X$  или прилагательное входят в состав сочинительных или близких к ним особых комитативных кон-

струкций: имеются в виду конструкции с так называемым "сочинительным сокращением" (см., например, [10], стр. 161 и сл.), то есть выражения типа а) замечательные = мн писатель X - ед и врач = ед, б) молчаливые [Y = мн] отец [X = ед] с сыном [Y = ед]; в) красный [Y = ед] и зеленый [Y = ед] шары [X = мн].

Нет сомнения, что две последние группы условий особого согласования X-а и Y -а представляют наибольший теоретический интерес. Природа и характер этих условий, как мы сейчас увидим, различны, и различны самым решительным образом.

Особое соответствие характеристик прилагательного Y и существительного X в "малых" количественных группах - это своего рода синтаксическая аномалия. Отсутствие согласования X-а и Y -а по числу (или по числу и падежу) объясняется здесь, как известно, историческими причинами: можно, по-видимому, считать, что в количественных группах типа два стола, четыре стены и т.п. стола и стены - реликт формы именительно-винительного падежа двойственного числа существительного, омонимичной форме родительного падежа единственного числа<sup>5</sup>, а у прилагательного форма двойственного числа не сохранилась [II, с. 207-215]. Это явление представляется - по крайней мере для современного состояния языка - достаточно случайным. Добавим, что такое согласование прилагательного с существительным - явление весьма локальное: значения характеристик здесь не совпадают только тогда, когда числительное стоит в именительном или (при неодушевленном X-е) винительном падежах (ср. два круглых стола - двумя круглыми столами). Этот факт целиком относится к (поверхностному) синтаксису и не несет никакой семантической нагрузки.

Что касается особого согласования прилагательного с существительным в сочинительных или комитативных группах, то этот факт представляется гораздо более глубоким и принципиальным.

Чтобы прояснить существо дела, рассмотрим конструкции с сочинительным сокращением и их возможные представления на поверхностном уровне несколько подробнее. (Для простоты мы ограничимся здесь сочинительными конструкциями с союзом и:

<sup>5</sup> Ср., однако, случаи, когда данная омонимия не имеет места: не прошел ни шага [= ед, род], но два шага [≠ ед, род], с этого часа, но четыре часа. У существительных шаг, час и некоторых других усматривается особая счетная форма [3, с. 302].

другие сочинительные конструкции и комитативные группы ведут себя похожим, хотя и не вполне тождественным образом).

Прежде всего отметим следующую важную особенность русских конструкций с сочинительным сокращением: число прилагательного и глагола - обычно чисто синтаксический компонент морфологической характеристики - может быть здесь семантически нагруженным; ср.

(1) Здесь работают замечательные писатель и врач (= '... замечательный писатель и замечательный врач');

(2) Здесь работают замечательный писатель и врач ( $\approx$  '... замечательный писатель и какой-то врач');

(3) Здесь работает замечательный писатель и врач (два значения: а) '... один человек - замечательный писатель и он же замечательный врач'; б) '... один человек - замечательный писатель и он же какой-то врач');

(4) Здесь работают замечательные писатели и врачи (четыре значения: а) '... замечательные писатели и замечательные врачи'; б) '... замечательные писатели и какие-то врачи'; в) '... замечательные писатели и они же замечательные врачи'; г) '... замечательные писатели и они же какие-то врачи');

(5) Имеются красные и зеленые шары ('... красные шары и зеленые шары');

(6) Имеются красный и зеленый шары ('...один красный шар и один зеленый шар').

Посмотрим теперь, каким должно быть поверхностно-синтаксическое представление для фраз типа (1)-(6). Все эти фразы различаются как по смыслу, так и формально; ясно поэтому, что и ПСП этих фраз должны быть различны. Между тем, по соглашению [9, с. 210], число прилагательных и глаголов, будучи синтаксически обусловленной характеристикой, в поверхностно-синтаксической структуре (ПСС), вообще говоря, отражаться не должно; если принять это соглашение, то окажется, что фразы (1), (2) и (3), а также фразы (5) и (6) получают одинаковые ПСС. Отсюда сразу же следует, что необходимо либо изменить вид ПСС для подобных выражений, либо иметь какие-то дополнительные средства их различия с помощью других компонентов ПСП. В этой связи предлагалось по крайней мере четыре решения:

1) отказаться от требования древесности ПСС и допускать синтаксическое подчинение одного узла целой группе узлов[4];

фрагменты ПСП для фраз (1) и (2) выглядели бы тогда, например, так:

ЗАМЕЧАТЕЛЬНЫЙ  $\xrightarrow{\text{опред}}$  ПИСАТЕЛЬ<sub>ед</sub>  $\xrightarrow{\text{сочин}}$  И  $\xrightarrow{\text{сочин-связн}}$  ВРАЧ<sub>ед</sub> (для 1)

ЗАМЕЧАТЕЛЬНЫЙ  $\xleftarrow{\text{опред}}$  ПИСАТЕЛЬ<sub>ед</sub>  $\xrightarrow{\text{сочин}}$  И  $\xrightarrow{\text{сочин-связн}}$  ВРАЧ<sub>ед</sub> (для 2)

2) ввести для конструкций с сочинительным сокращением особые ПСО [9, с. 214];

3) пользоваться в ПСП, дополнительно к дереву зависимостей, скобочной структурой [9, с. 215]; ср.

ЗАМЕЧАТЕЛЬНЫЙ  $\xleftarrow{\text{опред}}$  (ПИСАТЕЛЬ<sub>ед</sub>  $\xrightarrow{\text{сочин}}$  И  $\xrightarrow{\text{сочин-связн}}$  ВРАЧ<sub>ед</sub>) (для 1)

(ЗАМЕЧАТЕЛЬНЫЙ  $\xleftarrow{\text{опред}}$  ПИСАТЕЛЬ<sub>ед</sub>  $\xrightarrow{\text{сочин}}$  И  $\xrightarrow{\text{сочин-связн}}$  ВРАЧ<sub>ед</sub>) (для 2)

4) пользоваться в ПСП сведениями о кореферентности (= тождестве денотатов) именных лексем - узлов ПСС [9, с. 219]<sup>6</sup>; например, фрагмент ПСП для фразы (3) имел бы тогда следующий вид:

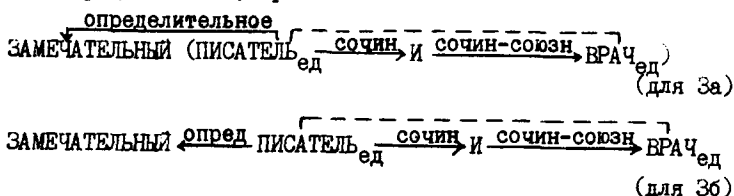
ЗАМЕЧАТЕЛЬНЫЙ  $\xleftarrow{\text{опред}}$  ПИСАТЕЛЬ<sub>ед</sub>  $\xrightarrow{\text{сочин}}$  И  $\xrightarrow{\text{сочин-связн}}$  ВРАЧ<sub>ед</sub>

Ни одно из этих решений, однако, не может спасти положения. В первых трех случаях (по-видимому, полностью эквивалентных) можно различить фразы (1) и (2), но нельзя отличить эти фразы от (3); в четвертом случае фраза (3) отличается от (1) и (2), но сами эти фразы остаются неразличимыми. Что касается фраз типа (5) и (6), которым разумно приписать одинаковую скобочную и анафорическую структуры, то они будут представлены тождественным образом во всех четырех случаях.

В качестве пятого решения можно было бы предложить следующее: попытаться использовать одновременно два из указанных средств, т.е., например, помечать в ПСП и скобочную структуру, и анафорические связи; в этом случае фразы типа (1), (2) и (3) будут различены. Однако это решение, по-види-

<sup>6</sup> Ср. [12], где сведения о кореферентности существенным образом используются в качестве особого компонента семантического представления фраз.

тому, еще менее удачно, чем предыдущие: в одних случаях оно оказывается чересчур сильным, порождая для фраз типа (3) два разных представления, ср.



а для фраз типа (4), как легко проверить, целых четыре представления; в других же случаях это решение является слишком слабым: разные фразы (5) и (6), как и при предыдущих решениях, будут представлены одинаково.

Имеется ли вообще сколько-нибудь естественный способ отразить в поверхностно-синтаксическом представлении те различия, которые передаются числом прилагательного и глагола в подобных конструкциях? По нашему мнению, такого способа нет и не может быть: механизмы, отвечающие здесь за выбор числа, не относятся к сфере поверхностного синтаксиса и должны формулироваться в терминах более глубокого - вероятнее всего, семантического - уровня. Действительно, правила, по которым язык выбирает число прилагательного или глагола в сочинительных конструкциях, совсем иные, нежели, скажем, правила выбора падежа прилагательного, и при этом гораздо более сложные; ср. колебания по числу во фразах, где падеж прилагательного бесспорен: Такая < такие > принципиальность и смелость проявилась < проявились? > в том, что ...; Мысль о независимости формы от субстанции и особенно вытекающее из нее положение о том, что язык может быть не звуковым, казалась < казались? > фантастической < фантастическими? > ; Советского < советских? > рабочего и колхозника такие проблемы не волнуют и т.д. (Многочисленные примеры такого рода колебаний см. в монографии [13]).

Учитывая сказанное выше, мы предлагаем следующее, весьма простое, решение: для всех сочинительных и подобных им конструкций удерживать в ПСС помету о числе прилагательного и глагола (и вообще: в тех случаях, когда это оправдано, использовать пометы о семантически наполненных морфологических характеристиках, даже если соответствующие характеристики обычно являются чисто синтаксическими: эти пометы должны учитываться - и самым существенным образом - на более глубо-

ких уровнях). Данное решение обладает по крайней мере следующими достоинствами: оно позволяет - в отличие от решений I-5 - получать разные ПСП для фраз с сочинительными конструкциями типа (I)-(6) и не порождает - в отличие от решений 4-5 - ложной поверхностно-синтаксической омонимии. Кроме того, наличие пометы о числе соответствует, на наш взгляд, содержательной стороне дела, позволяя отделить поверхностный механизм согласования прилагательного с существительным в сочинительных конструкциях по падежу и роду от сложного глубинного механизма (или ряда механизмов?) соответствия их числа.

Из принятого нами решения сразу же вытекает практическое следствие: поскольку число прилагательного в сочинительных и похожих конструкциях фигурирует как в ГМП, так и в ПСП и обеспечивать особое согласование по числу в этих конструкциях - вообще не дело поверхностного синтаксиса, то это согласование - в отличие от особого согласования в количественных группах - не имеет отношения к приводимому ниже оператору  $\text{СОГЛ}_{A(S)}$  и не должно в нем отмечаться.

Переходим теперь к формулировке правил оператора.

Оператор  $\text{СОГЛ}_{A(S)}$  ( $Y, X$ ) состоит из десятка правил, имеющих вид

$$X : Y \mid C$$

Ниже приводится семь основных правил этого оператора. В левой части каждого правила (до двоеточия) записываются морфологические характеристики и синтаксические особенности существительного  $X$ , отражающиеся на согласовании. В правой части правила (после двоеточия) приводятся необходимые морфологические характеристики прилагательного  $Y$ . Справа от вертикальной черты формулируются условия к правилу: требования к синтаксическому и линейному контексту для согласуемых словоформ.

В правилах I-7 используются следующие обозначения морфологических категорий, их значений и синтаксических признаков лексем и словоформ.

anim	: одушевленность
g	: род
n	: число
c	: падеж
муж	: мужской род
жен	: женский род

ед : единственное число  
 мн : множественное число  
 род : родительный падеж  
 парт : партитивный падеж  
 предл : предложный падеж  
 местн : местный падеж  
 колич : существительные типа миллион (см. правила I и 5)  
 притяж : притяжательные прилагательные мой, твой, ваш...  
 проф : существительные типа врач (см. правила I и 5)  
 цифр-I : цифровые обозначения числительных на I, но не на II (т.е. I, 2I, 3I, ..., IOI, I2I, ...);  
 цифр-мал: цифровые обозначения числительных, оканчивающиеся на 2, 3, 4, но не на I2, I3, I4 (т.е. 2, 3, 4, 22, 23, 24, ... IO2, IO3, IO4, ...).

В записи правил  $\bar{S}$  обозначает существительное, А - адъектив, т.е. прилагательное или причастие. Синтаксические (словарные) свойства слов записываются в скобках, морфологические характеристики вне скобок.

Все правила снабжены неформальными комментариями и примерами.

Оператор согласования СОГЛ<sub>A(S)</sub>, как и другие операторы, действующие в нашей системе поверхностного синтаксиса, построен дизъюнктивно: наличие согласования между двумя словами (в данном синтаксическом или линейном контексте) констатируется в том случае, если выполняются все условия хотя бы одного из правил оператора; в противном случае делается вывод об отсутствии согласования.

I. Согласование прилагательного с существительным в единственном числе без осложняющих количественных и сочинительных зависимых

$X = (S, g, \text{anim}), \text{ед}, c; Y = (A), g, \text{anim}, \text{ед}, c$

- 1) не X сочин  $\rightarrow$  Z
- 2) не X атриб  $\rightarrow$  Z<sub>с3</sub>
- 3) если X колич  $\rightarrow$  Z  
     то Z = ОДИН или  
         Z = (цифр-I)
- 4) если c  $\neq$  парт/местн,  
     то c' = c
- 5) если c = парт,  
     то c' = род
- 6) если c = местн,  
     то c' = предл

- 7)  $X \neq$  (проф)  
8)  $X \neq$  (колич)

Комментарии. I. Пометы  $g$  и  $an\bar{u}$  в обеих частях правила обеспечивают совпадение значений соответствующих грамматических категорий рода и одушевленности у  $X$ -а и  $Y$ -а (классифицирующих у существительного, словоизменительных у прилагательного)<sup>7</sup>.

2. Условия 1 и 2 запрещают наличие у существительного сочинительных или комитативных (вводимых предлогом  $C$  в особом, третьем, значении) зависимых: как мы видели, наличие подобных зависимых влияет на статус категории числа прилагательного и переводит его в ранг семантически содержательной характеристики.

3. Условие 3 накладывает ограничение на количественные зависимые существительного, допуская только такие из них, которые не влияют на согласование (а именно, числительное один, а также составное числительное, оканчивающееся на один, возможно записанное цифрами - такие числительные задаются посредством синтаксического признака "цифр-1").

4. Условия 4-6 задают соответствие падежей  $X$ -а и  $Y$ -а: если падеж существительного - не партитивный и не местный, то падеж  $Y$ -а ( $c'$ ) совпадает с падежом  $X$ -а ( $c$ ); если падеж  $X$ -а - партитивный, то  $Y$  стоит в родительном падеже, а при  $X$ -е в местном падеже требуется предложный падеж  $Y$ -а.

5. Условие 7 требует, чтобы существительное  $X$  не было словом типа врач, допускающего колебания в числе согласуемого с ним определения, а условие 8 - чтобы существительное не было словом типа миллион, допускающего колебания в числе согласуемого с ним определения (см. об этом выше). Эти случаи описываются особыми правилами оператора, которые, однако, здесь не приводятся вследствие явной периферийности.

Примеры. (1) большой [ $Y$ ] квадрат [ $X$ ]; (2) графика [ $X_1$ ], соответствующего [ $Y_1$ ] данной [ $Y_2$ ] функции [ $X_2$ ]; (3) пятью [ $Y$ ] зеленую [ $Y_2$ ] ракету [ $X_1, X_2$ ]; (4) день [ $X$ ] на седьмой [ $Y$ ]; (5) заметил высокого [ $Y$  = вин. одуш.] негра [ $X$  = (одуш.), вин]; (6) Он известный [ $Y$ ] зануда [ $X$  = (муж)]; (7) Он <она> известная [ $Y$ ] зануда [ $X$  = (жен)] (см. примечание); (8) одно < двадцать одно > каменное [ $Y$ ] здание [ $X$ ]; (9) купил год-

<sup>7</sup> О классифицирующих и словоизменительных грамматических категориях см. [5].



ландского [Y = род] сыру [X = парт]; (IO) в глубоком [Y = = предл] тылу [X = местн].

Примечание. Для слов так называемого "общего рода" типа зануда, умница, староста и т.п., допускающих колебания в согласовании по роду, предлагается следующее словарное описание:

зануда I (женского рода)  $\approx$  'нудный человек';

зануда 2 (мужского рода)  $\approx$  'нудный мужчина'.

Тогда в выражениях типа Он известная зануда и Она известная зануда реализуется первое значение существительного, а в выражениях типа Он известный зануда - второе значение. NB: в сочетаниях типа \*Она известный зануда, на наш взгляд, нарушено не синтаксическое, а семантическое согласование; в самом деле, такое сочетание должно было бы означать примерно следующее: 'Она, как известно, весьма нудный мужчина', - однако этот в высшей степени нестандартный смысл должен выражаться более эксплицитным образом, например, прямо так: Она, как известно, нудный мужчина. (О семантическом согласовании см. выше).

Не все слова общего рода устроены одинаковым образом; так, для слов староста и сирота требуются другие описания:

староста I (мужского рода)  $\approx$  'руководитель';

староста 2 (женского рода)  $\approx$  'руководитель-женщина'

(ср. Он - наш новый староста, Она - наш новый <наша новая> староста, но не \*Он - наша новая староста);

сирота I (мужского рода)  $\approx$  'мужчина, не имеющий отца или матери'

сирота 2 (женского рода)  $\approx$  'женщина, не имеющая отца или матери'

(ср. Он - круглый сирота, Она круглая сирота, но не \*Он - круглая сирота или \*Она круглый сирота).

Согласование прилагательного с существительным в единственном числе при осложняющем количественном зависимом (2-4)<sup>8</sup>

<sup>8</sup> Правила 2-4 описывают согласование прилагательного с существительным не для любых малых количественных групп, а лишь для таких, в которых существительное имеет форму единственного числа (и родительного падежа: последний не отмечается в правилах оператора по той причине, что другого падежа здесь не может быть). Согласование прилагательного с существительным в других малых количественных группах, в том числе во всех косвенных (двумя большими глотками, двух больших глотков) и в тех прямых, где су-

## 2. Обычное прилагательное + существительное мужского или среднего рода

$X = (S, \underline{\text{не жен}}, \text{ед} : Y = (A),$ <div style="text-align: right;">мн, род</div>	1) $X \xrightarrow{\text{колич}} Z$
	2) $Z = (\text{мал})$ или $Z = (\text{цифр-мал})$
	3) $Z \neq (\text{притяж})/\text{ЭТОТ}$

Комментарии. 1. Условия 1 и 2 требуют наличия при X-е малого числительного, определяющего для X-а и Y-а особое согласование.

2. Условие 3 отвергает некоторые местоименные прилагательные, допускающие два варианта согласования с существительным в малых количественных группах. Эти варианты согласования описываются правилом 3.

Примеры. (1) два [Z] больших [Y] кирпичных [Y<sub>2</sub>] здания [X<sub>1</sub>, X<sub>2</sub>]; (2) оба [Z] придаточных [Y] предложения X; (3) полтора [Z] свежих [Y] бифштекса X; (4) три [Z] юных [Y] пажа [X]; (5) "Сто четыре [Z] новых [Y] тома [X] там стояли под стеклом" (С.Я. Маршак); (6) 33 [Z = (цифр-мал)] центнера [X] ржи, поступивших [Y] вчера... : в конструкциях типа 33 центнера, поступившие вчера, указанные 42 правила и т.п. мы считаем прилагательное зависящим от числительного; по этой причине такие конструкции на описываются данным оператором.

## 3. Особое местоименное прилагательное + существительное мужского или среднего рода

$X = (S, \underline{\text{не жен}}, \text{ед} : Y = (A) \text{ мн, род}$ <div style="text-align: right;"><u>или</u> (A) мн, с</div>	1) $X \xrightarrow{\text{колич}} Z$
	2) $Z = (\text{мал})$ или $Z = (\text{цифр-мал})$
	3) $Y = (\text{притяж})$ или $Y = \text{ЭТОТ}$
	4) $c(Y) = c(Z)$

Комментарии. 1. Об условиях 1-3 см. комментарии к правилу 2.

2. Условие 4 устанавливает соответствие падежей числительного и прилагательного во втором варианте согласования

существительное - адъективного склонения (три московские булочные, оба относительных прилагательных) описываются правилом 6 - общим правилом согласования прилагательного с существительным множественного числа.

(вижу оба [= вин] твои [= вин] ключа; оба [= им] твои = им ключа лежат здесь ...).

Примеры. (1) 3 [Z] ваши < ваших > [Y] предложения [X]; (2) оба [Z] наши < наших > [Y] алгоритма [X]; (3) 24 эти < этих > [Y] явления [X].

#### 4. Прилагательное + существительное женского рода

X=(S, жен), ед : Y=(A) мн, с  
или (A) мн, род

1) X <u>колич</u> → [Z]
2) Z = (мал) или Z = (цифр-мал)
3) с (Y) = с (Z)

Об условиях см. комментарии к правилам 2 и 3

Примеры. (1) получили две [Z = вин] большие [Y = вин] больших [Y = род] комнаты [X]; (2) три [Z] красные < красных > [Y] груши [X]; (3) две [Z] непрерывные < непрерывных > [Y] функции [X] и график; (4) обе [Z] ваши < ваших > [Y] записки [X].

Примечание. В правилах 3 и 4 мы позволили себе для простоты отвлечься от тонких и трудноуловимых семантических различий между вариантами согласования с родительным и именительным/винительным падежами (ср. Он съел три большие груши - скорее, никаких других груш не было - и Он съел три больших груши - вероятно, были и другие груши). Если учитывать такие различия, то падеж прилагательного в подобных сочетаниях окажется значащим; в этом случае с ним следовало бы поступать так же, как с числом прилагательного в сочинительных конструкциях, т.е. обрабатывать на более глубоком, чем поверхностно-синтаксический, уровне представления.

#### 5. Согласование прилагательного с существительным в единственном числе, имеющим сочинительные или комитативные зависимые

X=(S, g, anim), ед, с : Y=(A), g<sup>9</sup>,  
anim, ед, с

1) X <u>сочин</u> → Z <u>или</u> X <u>атриб</u> Z c <sub>3</sub>
2) <u>если</u> X <u>колич</u> → w, <u>то</u> w ≠ (мал) и

- $w \neq$  (цифр-мал)  
 3) если  $c \neq$  парт/местн,  
 то  $c' = c$   
 4) если  $c =$  парт, то  
 $c' =$  род  
 5) если  $c =$  местн, то  
 $c' =$  предл  
 6)  $X \neq$  (проф)  
 7)  $X \neq$  (колич)

Комментарии. I. 0 пометах  $g$  и  $anim$ , условиях 3-7 см. комментарии к правилу I.

2. Условие I требует наличия при X-е сочинительных или комитативных зависимых. Как было показано выше, число прилагательного в таких конструкциях приобретает семантическую нагрузку и обрабатывается отдельно от оператора  $COGL_A(S)$ .

3. Условие 2 отвергает малые количественные зависимые при X-е: конструкции типа две непрерывных функции и график, 24 просторных класса и кабинета и пр. описываются правилами 2-4.

Примеры. (1) видели пожилого [Y] мужчину [X] и [Z] женщину - видели пожилых [Y] мужчину [X] и [Z] женщину; (2) прямая X и [Z] окружность < окружности >, изображенные [Y] на экране; (3) первый [Y] оператор [X], или [Z] преобразователь; (4) расположенными [Y<sub>1</sub>] в третьем столбце приближенной [Y<sub>2</sub>] формулой [X<sub>1</sub>, X<sub>2</sub>], символом [Z] Кроникера и неопределенным интегралом; (5) прибывшая < прибывшие > [Y] вчера женщина [X] с [Z] ребенком; (6) хороший [Y] работник [X] и [Z] семьянин; (7) А.П. Чехов, замечательный [Y<sub>1</sub>] русский [Y<sub>2</sub>] писатель [X<sub>1</sub>, X<sub>2</sub>] и врач ...; (8) отличаются высоким [Y<sub>1</sub>] профессиональным [Y<sub>2</sub>] уровнем [X<sub>1</sub>, X<sub>2</sub>] и [Z] подготовкой; (9) для простого [Y] рабочего X или [Z] крестьянина.

6. Согласование прилагательного с существительным  
 во множественном числе (6-7)

6. Прилагательное не имеет сочинительных зависимых

$X = (S, anim), c : Y = (A), anim, c \quad | I) \text{ не } Y \xrightarrow{\text{сочин}} Z$

<sup>9</sup> Указание о роде Y в правой части правила 5 (а также правила 7, см. ниже) оказывается релевантным (= приписывается при синтезе и элиминируется при анализе) тогда, когда Y - в единственном числе.

Комментарий. Условие I отвергает конструкции типа красный и зеленый шары, красный и зеленые шары, красные и зеленые шары и т.п. Эти конструкции описываются правилом 7.

Примеры. (1) большие [Y] квадраты [X]; (2) каменными [Y] зданиями [X]; (3) трем вашим [Y] дочерям [X]; (4) два <двое> русских [Y] рядовых [X]; пять окон [X], открытых [Y] настель; (6) встретил молодых [Y = вин. одуш] негров [X = (одуш), вин]; (7) замечательные [Y] писатели [X] и врачи <врач> .

7. Прилагательное имеет сочинительное зависимое

$$X = (S, g, \text{anim}), \text{мн}, c : Y = (A), g, \text{anim}, c \mid I) Y \xrightarrow{\text{сочин}} Z$$

Комментарий. Условие I требует наличия у Y -а сочинительных зависимых: мы видим, что в таком случае число прилагательного становится значащим и не может обеспечиваться оператором  $\text{СОГЛ}_A(S)$  Ср. также правило 5 и комментарий к нему.

Примеры. (1) первый [Y] и [Z] второй случаи [X] рассматривались ...; (2) красными [Y] и [Z] зелеными шарами [X]; (3) ни того [Y], ни [Z] другого вариантов [X]; (4) стальной [Y] или [Z] никелевые резцы [X] ...

Примечание. Конструкции типа два больших и маленький шары, маленькие и 2 больших шара, 2 больших и 5 маленьких шаров и т.п. признаются эллиптическими и не описываются оператором  $\text{СОГЛ}_A(S)$ .

## Литература

1. Апресян Ю.Д. Языковая аномалия и логическое противоречие. - В кн.: *Tekst. Język. Poetyka*. Wrocław, Oszolineum, 1978.
2. Бабицкий К.И. Согласование членов как необходимое условие правильности предложения. - Информационные вопросы семиотики, лингвистики и автоматического перевода. Вып. 4. М., 1974, с. 63-76.
3. Виноградов В.В. Русский язык (грамматическое учение о слове). М.-Л., Учпедгиз, 1947.
4. Гладкий А.В. Об описании синтаксической структуры предложения. - *Computational Linguistics*, VII, 1968, с.21-44.
5. Зализняк А.А. Русское именное словоизменение. М., "Наука", 1967.

6. Иомдин Л.Л. Некоторые трудные вопросы русского согласования. - В кн.: Семантический и синтаксический компоненты лингвистического обеспечения. Новосибирск, 1979.
7. Иомдин Л.Л., Мельчук И.А., Перцов Н.В. Фрагмент модели русского поверхностного синтаксиса. I. Предикативные синтагмы. - НТИ, сер II, 1975, № 7, с. 30-43.
8. Иомдин Л.Л., Перцов Н.В. Фрагмент модели русского поверхностного синтаксиса. II. Комплективные и присвязочные конструкции. - НТИ, сер. II, 1975, № II, с. 22-32.
9. Мельчук И.А. Опыт теории лингвистических моделей "Смысл  $\longleftrightarrow$  Текст". М., "Наука", 1974.
10. Падучева Е.В. О семантике синтаксиса. М., "Наука", 1974.
11. Шахматов А.А. Историческая морфология русского языка. М., Учпедгиз, 1957.
12. Jackendoff, R.S. Semantic Interpretation in Generative Grammar. Cambridge (Mass.) - London, MIT Press, 1972.
13. Crockett, D.B. Agreement in Contemporary Standard Russian. Cambridge (Mass.) - Slavica Publishers, 1976.

ON WORKING OUT A FORMAL MODEL OF MODERN RUSSIAN:  
AN OPERATOR OF SYNTACTIC AGREEMENT  
FOR ATTRIBUTIVE CONSTRUCTIONS

L. Yomdin  
S u m m a r y

Every artificial intelligence system of any degree of sophistication should incorporate a reliable working model of natural language with a full-fledged syntactic component.

The operator of syntactic agreement for Russian attributive constructions, suggested in the present paper, is part of such a component. It gives a formalized description of the principal type of Russian agreement, namely, the agreement between an adjective and a noun in nominal phrases like *vysokaya sosna* (a high fir-tree).

Based on the principles of the "Meaning-Text" linguistic theory, the agreement operator may be considered as a separate fragment of a formal model of Russian. It consists of a number of rules all ready for computer implementation and may be used as an independent mechanism in the automatic syntactic analysis and synthesis of Russian texts.

## АЛГОРИТМЫ ПЕРЕВОДА С $\Psi$ -ЯЗЫКА НА РУССКИЙ

М. Койт

### Введение

Проблема использования естественного языка в системах искусственного интеллекта имеет три основных стороны [13]:

- 1) перевод исходного текста в некоторое представление его смысла;
- 2) обработка смыслов, сохраняемых в памяти системы, на основании воспринятого сообщения;
- 3) обратный переход от представления смысла к тексту на естественном языке.

В настоящее время работы по созданию диалоговых систем на языке, близком к естественному, получают все большее развитие [1, 2, 8, 12-14]. В разрабатываемых отечественных вопросно-ответных системах для представления смыслов используется предпочтительно исчисление предикатов или реляционный язык [5, 10].

Выдающаяся вопросно-ответная система ДИЛОС, разрабатываемая в Вычислительном центре АН СССР, представляет собой совокупность ЛИСП-программ для ЭВМ БЭСМ-6 [1]. Информация от пользователя поступает на вход системы в виде печатного текста на обычном русском (или каком-нибудь другом) языке. Эта информация поступает в лингвистический процессор, где она переводится [11] на внутренний язык системы -  $\Psi$ -язык, который является реляционным языком. В зависимости от типа сообщения на  $\Psi$ -языке, анализатор посылает его в те или иные исполнительные процессоры (информационный, логический, вычислительный). Система ДИЛОС может быть применена для решения информационных, вычислительных и логических задач и для осуществления диалога между человеком и ЭВМ на языке, максимально приближенном к естественному [9]. Ответ пользователю выдается в виде  $\Psi$ -выражения. Если это выражение имеет простую структуру, то пользователь его, наверное, понимает. В противном случае, однако, возникает проблема перевода с  $\Psi$ -языка на естественный, которой до сих пор уделено меньше внимания, чем остальным аспектам использования естественного языка.



В связи с этим перед нами поставлена задача перевода с  $\varphi$ -языка на обычный русский язык. В настоящей статье описываются два близких друг к другу алгоритма такого перевода. По общему принципу излагаемые здесь алгоритмы развивают идеи, лежащие в основе семантического алгоритма, описанного нами в статье [6].

## 1. Описание первого алгоритма

### 1.1. Разложение $\varphi$ -выражения на простые части

Мы исходим из определения  $\varphi$ -выражения, данного в статье [7]. Алгоритм строится в предположении, что заданные отношения -  $n_i$ -арные.

Обработка исходного  $\varphi$ -выражения начинается с его разложения на совокупность простых скобочных выражений. Для этого  $\varphi$ -выражение анализируется справа налево. Первое справа вхождение простого реляционного выражения вида запоминается

$$(a_{i_1} r_j a_{i_2} \dots a_{i_{n_j}}) \quad (1)$$

как первый член образуемого разложения и заменяется в анализируемом  $\varphi$ -выражении его главным аргументом  $a_{i_1}$ , перед которым ставится метка  $x$ . Дальше продолжается анализ выражения аналогичным образом: постепенно выделяются подвыражения вида (1) (возможно, содержащие уже вхождения метки  $x$ ) и соответственно сокращается выражение. Если очередное (справа) простое скобочное выражение является кванторной частью

$$(X_j a_i) \quad (2)$$

$\varphi$ -выражения, то ее в качестве очередного члена включают в образуемое разложение, а в анализируемом выражении она заменяется меткой  $x$ . Описанный процесс кончается, если исходное  $\varphi$ -выражение сокращено до цепочки, состоящей только из понятий  $a_i$  и меток  $x$ . Тем самым образована совокупность

$$M_1, M_2, \dots, M_p \quad (3)$$

простых скобочных выражений. Процесс образования разложения (3) можно представить как последовательное отсекание дерева, соответствующего исходному  $\varphi$ -выражению, происходящее снизу вверх и справа налево. Отсеченные поддеревья и образуют совокупность (3).

Пример 1. Для  $\varphi$ -выражения

$$(a_1 r_2 ((a_3 r_3 a_4) r_1 (a_2 r_3 a_4)))$$

получаем разложение  $M_1 = (a_2 r_3 a_4)$ ,  $M_2 = (a_3 r_3 a_4)$ ,

$$M_3 = (a_3 r_1 a_2), \quad M_4 = (a_1 r_2 \times a_3).$$

## 1.2. Синтез текста

### 1.2.1. Исходные данные

По найденному разложению будем синтезировать текст. Для образования текста нужно знать:

1) тексты понятий  $a_i \in A$ ;

2) предписания, по которым образуются тексты простых скобочных выражений вида (1) или (2) из текстов понятий, в них входящих.

Сделаем предположение, что текстом любого понятия является некоторая словоформа. Поскольку текст понятия в сложное  $\varphi$ -выражение может входить в измененном виде (другой падеж, число и т.п.), оказывается целесообразным задавать не сами тексты  $T a_i$  понятий (словоформы), но т.н. сырые тексты  $T^0 a_i$  - совокупности значений информационных признаков, по которым позднее можно образовать словоформы. Для представления сырых текстов можно выделить информационные признаки, приведенные в таблице 1.

Т а б л и ц а 1

П р и з н а к		З н а ч е н и я
1	:	2
<основа>		перечислены в заданном словаре
<часть речи>		Сущ. (имя существительное), Гл (глагол) и т.д.
<род>		Муж, Жен, Ср
<число>		Ед, Мн
<падеж>		Им, Род, Дат, Твор, Предл
<время>		Наст, Прош, Буд
<наклонение>		Изъяв, Сослаг, Повелит
<вид>		Сов, Несов
<префикс>		перечислены в заданном словаре
<лицо>		1, 2, 3
<частица>		[ся], [сь]
<отрицание>		[не]

Если некоторый признак отсутствует, то его значением можно читать пустое слово  $\Lambda$ .

Сырые тексты понятий могут быть заданы следующими совокупностями. Для измененных слов:

$\langle \text{основа} \rangle$ ,

для склоняемых слов:

$(\langle \text{основа} \rangle \quad \langle \text{часть речи} \rangle \quad \langle \text{род} \rangle \quad \langle \text{число} \rangle \quad \langle \text{падеж} \rangle)$ ,

для глаголов:

$(\langle \text{отрицание} \rangle \quad \langle \text{префикс} \rangle \quad \langle \text{основа} \rangle \quad \langle \text{часть речи} \rangle \quad \langle \text{наклонение} \rangle \quad \langle \text{вид} \rangle \quad \langle \text{время} \rangle \quad \langle \text{род} \rangle \quad \langle \text{число} \rangle \quad \langle \text{частица} \rangle)$ .

Предписания для образования текстов простых скобочных выражений также задаем в виде сырых текстов соответствующих выражений. Сделаем предположение, что текст любого выражения (I) является линейной функцией от функций текстов его аргументов:

$$T(a_{i_1} r_j a_{i_2} \dots a_{i_n}) = \varphi^j_1(Ta_{i_1}), \varphi^j_2(Ta_{i_2}), \dots, \varphi^j_{n_j}(Ta_{i_{n_j}}) = \\ = \varphi^j_{k_1}(Ta_{i_{k_1}}) \varphi^j_{k_2}(Ta_{i_{k_2}}) \dots \varphi^j_{k_{n_j}}(Ta_{i_{k_{n_j}}}) \varphi^j_{n_j+1},$$

где  $\varphi^j_{k_i}$  - цепочки словоформ или знаков препинания;

$$\{k_1, \dots, k_{n_j}\} = \{1, \dots, n_j\}.$$

Аналогично, для выражения (2) пусть

$$T(\chi_j a_i) = \chi^j(\psi^j(a_i)) = h^j_1 \psi^j(a_i) h^j_2$$

(об использовании кванторов см. [3-4]).

При синтезе текста придется учитывать и грамматическое согласование: изменение значений информационных признаков одной (главной) словоформы в образуемом тексте может обусловить необходимость к заменам и в других (второстепенных) словоформах. Согласование указываем меткой  $\star$  перед изменяемыми признаками во второстепенных словоформах. Если порядок информационных признаков строго зафиксирован, то значение признака, подлежащего согласованию, можно опускать, поскольку его можно определить по главной словоформе.

Пример 2. Сырые тексты простых реляционных выражений могут быть заданы следующим образом:

$$T^0(a_1 r_2 a_3) = (\Lambda \text{ I}([Erop] \text{ Сущ Муж Ед Им}) \\ \varphi^2_1(T^0 a_1) = T^0 a_1$$

( $\Lambda$   $\Lambda$  [двига] Гл Изъяв Несов Наст  $\&$ Муж  $\&$ Ед  $\&$ З [ся]) [к]  
 2([озер] Сущ Ср Ед Дат)  $\Lambda$  )

$$\varphi^2_x(T^0a_3)$$

(текст: Егор двигается к озеру);

$$T^0(a_2 \ \&_3 \ a_4) = (\Lambda \ 2([больш] Прил \&Муж \&Ед \&Им) \Lambda$$

$$\varphi^3_x(T^0a_4) = T^0a_4$$

$$I([лес] Сущ Муж Ед Им) \Lambda )$$

$$\varphi^3_1(T^0a_2) = T^0a_2$$

(текст: большой лес)

и т.д.

## 1.2.2. Синтаксический этап синтеза

Синтез текста выполняется в два этапа. На первом (синтаксическом) этапе синтезируется сырой текст - цепочка совокупностей значений информационных признаков. На втором (морфологическом) этапе образуются словоформы.

Синтаксический синтез происходит по образованному разложению (3) начиная с последнего члена  $M_p$ . Это соответствует перемещению сверху вниз и слева направо по дереву исходного  $\varphi$ -выражения. Если  $M_p$  не единственный член в разложении (исходное выражение не является простым выражением вида (1)), то он обязательно содержит хотя бы одно вхождение метки  $x$ . Находим номер  $k$  аргументу (понятия  $a_i$ ), перед которым стоит первое слева вхождение метки, и по номеру  $k$  перенесем эту метку из  $M_p$  в заданный сырой текст  $T^0(M_p)$ , к началу подтекста  $\varphi^q_r(T^0a_i)$  соответствующего понятия. Получаем текст : $r$ : ( $T^{0x}(M_p)$ ), где метка : $r$ : указывает на номер члена разложения. На следующем шаге отмеченный подтекст  $\varphi^q_r(T^0a_i)$  приходится заменить текстом  $T^0(M_j)$  следующего члена  $M_j = M_{p-1}$  разложения (3), где  $\varphi^q_r(T^0a_i)$  является текстом главного аргумента (по дереву это будет сдвиг вниз). Перед заменяющим текстом  $T^{0x}(M_j)$  ставим метку : $j$ :. Если у текста  $\varphi^q_r(T^0a_i)$  значения некоторых информационных признаков отличаются от значений соответствующих информационных признаков в  $\varphi^{q-1}_r(T^0a_i)$ , то введем изменения в сырой текст  $T^{0x}(M_j)$ , учитывая по меткам  $\&$  и согласование. Аналогичным способом продолжаем перемещение по дереву вниз и постепенное дополнение сырого текста до тех пор, пока не доходим до такого члена  $M_j$  разложения, в кото-

рый не входит метка  $x$ . Это означает, что все понятия, являющиеся аргументами  $M_j$ , в исходном дереве  $\varphi$ -выражения оказываются висячими вершинами. Если  $j = 1$ , то синтез сырого текста этим окончен. В противном случае придется подниматься по дереву, пока не удастся перейти к первой слева еще не пройденной ветви. В разложении соответствующий член  $M_h$  ( $j < h \leq p$ ) содержит хотя бы одно вхождение метки  $x$ . В члены  $M_j, \dots, M_{h-1}$  метка  $x$  уже не входит. Находим номер  $k$  аргумента  $a_i$ , перед которым стоит первое слева вхождение метки  $x$ , и переносим метку от этого аргумента в синтезируемый текст, к сырому тексту  $\varphi^h_k(T^0 a_i)$ , который можно найти по метке  $h$  и номеру  $k$ . Далее продолжаем синтез вышеописанным способом. Если ни один из членов  $M_h$  ( $j \leq h \leq p$ ) не содержит метки  $x$ , то исходному  $\varphi$ -выражению соответствует более, чем одно дерево (лес), и со следующего члена  $M_{j-1}$  разложения начинается синтез сырого текста по следующему дереву.

Пример 3. Рассматриваем  $\varphi$ -выражение примера 1. Пусть заданы сырые тексты  $T^0 a_1, T^0 a_2, T^0 a_3, T^0 a_4, T^0(a_1 \tau_2 a_3), T^0(a_2 \tau_3 a_4), T^0(a_3 \tau_3 a_4)$  (ср. пример 2) и

$$\bullet \quad T^0(a_3 \tau_1 a_2) = (\bigwedge I([озер] \text{ Сущ Ср Ед Им}) \\ \varphi'_1(T^0 a_3) = T^0 a_3$$

$$[, \text{которое}] (\bigwedge \bigwedge [\text{наход}] \text{ Гл Изъяв Несов Наст } \times \text{Ср} \times \text{Ед} \\ \times \text{З} [\text{ся}]) [\text{в}] 2([\text{лес}] \text{ Сущ Муж Ед Предл}) \bigwedge ) \\ \varphi'_2(T^0 a_2)$$

(текст: озеро, которое находится в лесу).

Синтезируемый сырой текст по разложению, найденному в примере 1. Опуская, в целях лучшей обзорности, все информационные признаки, значения которых легко можно восстановить по начальным данным, последовательно получаем:

$$1) T^{0x}(M_4) = T^{0x}(a_1 \tau_2 a_3) = :4:(I([Егор] \dots) \\ ([\text{двига}] \dots) [к] 2 \times ([озер] \dots \text{Дат})), \\ M_4 = (a_1 \tau_2 a_3).$$

Теперь сравниваем отмеченный подтекст ( $[озер] \text{ Сущ Ср Ед Дат}$ ) с текстом ( $[озер] \text{ Сущ Ср Ед Им}$ ) главного аргумента реляционного выражения ( $a_3 \tau_3 a_2$ ). Обнаружив разницу в значениях последнего информационного признака, заменим в подтексте

$\varphi_2^2(T^0 a_3)$  значение Им на Дат и просмотрим весь текст  $T^0(M_3)$  с целью выяснения согласования. Остальные значения информационных признаков не подлежат замене. Подставляя текст  $T^{0X}(M_3)$  в синтезируемый, получаем:

2) :4: (I ([Егор]) ([двигā]) [к] 2:3: (I x ([озер] ...  
... Дат) [, которое] ([наход]) [в] 2( [лес] ... Предл) ));  
 $M_3 = (a_3 \tau_3 x a_2)$ .

Сравниваем новый отмеченный подтекст ([озер] ... Дат) с текстом ([озер]... Им) главного аргумента реляционного выражения  $M_2 \doteq (a_3 \tau_3 a_4)$ . Учитывая и согласование, получаем новый текст:

3) :4: (I ([Егор]) ([двигā]) [к] 2:3: (I:2:(2 ([больш] ... иДат) I ([озер] ... Дат) [, которое] ([наход]) 2([лес] ... Предл) ))).

Поскольку  $M_2$  не содержит метки x, возвращаемся к  $M_3$ , аргумент с номером 2 которого снабжен меткой x. По метке :3: и номеру 2 перенесем метку x в образуемой текст:

4) :4: (I ([Егор]) ([двигā]) [к] 2:3: (I:2:(2 ([больш] ... иДат) I ([озер] ... Дат) [, которое] ([наход]) [в] 2 x ([лес] ... Предл) )));

$M_3 = (a_3 \tau_1 a_2)$ .

Наконец, присоединяем к синтезируемому сыроу текст  $T^0(M_1) = T^0(a_2 \tau_3 a_4)$ , до этого выполнив согласование в значениях последнего информационного признака по имеющемуся сырому тексту ([лес] ... Предл):

5) :4: (I ([Егор]) ([двигā]) [к] 2:3: (I:2:(2 ([больш] ... иДат) I ([озер] ... Дат) ) [, которое] ([наход]) [в] :I: (2 ([больш] ... иПредл) I ([лес] ... Предл) ))).

### 1.2.3. Морфологический этап синтеза

Последний этап синтеза - морфологический - предполагает наличие словарей основ и списков окончаний. При синтезе произвольной словоформы по номеру основы находится соответствующая основа в словаре. Далее обращаемся к некоторому списку окончаний. Идентификатор списка дается в виде совокупности информационных признаков:

⟨часть речи⟩ ⟨род⟩ ⟨число⟩

(для падежных окончаний) или

⟨часть речи⟩ ⟨вид⟩ ⟨время⟩ ⟨род⟩ ⟨число⟩

(для отглагольных окончаний). Требуемое окончание в списке разыскивается по значению информационного признака ⟨падеж⟩

или, соответственно, <лицо> . Нужные формы сослагательного или повелительного наклонения образуются при помощи форм изъявительного наклонения. Поскольку имеется несколько типов склоняемых слов и глаголов, представляется целесообразным брать в качестве значений информационного признака <часть речи> не традиционные названия частей речи, а названия (номера) некоторых более мелких, подходящим образом выбранных подклассов существительных, глаголов и т.д.

Пример 4. Предполагая заданными нужные списки и словарь, получаем следующий результат морфологического синтеза сырого текста, образованного в примере 3:

Егор движется к большому озеру, которое находится в большом лесу.

## 2. О другом алгоритме перевода

Пусть заданы сырые тексты понятий и простых скобочных выражений, как это описано в п. 1.2.1. Анализируем исходное  $\varphi$ -выражение справа налево. Первое простое реляционное выражение вида (I) заменяем на его главный аргумент  $a_i$ , и запоминаем сырой текст опущенного выражения. Перемещаемся налево. Если находим выражение вида (I), не зависимое от выражений, опущенных на предыдущих шагах (ни один из его аргументов не является главным аргументом уже опущенного выражения), то опять выполняем сокращение найденного выражения до его главного аргумента и одновременно запоминаем его сырой текст, не зависимо от уже образованного текста. Если, однако, найденное выражение содержит один или несколько аргументов, представляющих более сложные  $\varphi$ -выражения, опущенные в ходе предыдущего анализа, то текст, который мы образуем по опускаемому выражению, получается из его сырого текста путем замены сырых текстов некоторых понятий (аргументов) уже синтезированными сырыми текстами. При этом приходится учитывать грамматическое согласование: если подтекст главного аргумента в заменяющем тексте имеет не те значения информационных признаков, что заменяемый текст понятия, то приходится вводить изменения в весь заменяющий текст. Следовательно, в таком случае необходимо снова просмотреть уже образованный кусок текста. Кванторные выражения, обнаруженные при перемещении налево, мы просто опускаем, а соответствующий текст присоединяем слева к тексту, образованному на предыдущем шаге. Синтез текста завершается морфологическим этапом, как это

описано в п. 1.2.3.

Иллюстрируем на примере применение второго алгоритма.

Пример 5. Исходим из  $\psi$ -выражения примера 1. Пусть заданы сырые тексты понятий и простых скобочных выражений, приведенные в примере 3. Применяя алгоритм, постепенно получаем следующие сырые тексты:

1) текст

$$T_1 = (2 ( [\text{больш}] \dots \text{жИм}) \text{ I } ([\text{лес}] \dots \text{Им}) );$$

сокращенное  $\psi$ -выражение

$$(a_1 \text{ } \text{ }_2 ((a_3 \text{ } \text{ }_3 a_4) \text{ }_1 a_2));$$

$$2) T_2 = (2 ( [\text{больш}] \dots \text{жИм}) \text{ I } ([\text{озер}] \dots \text{Им}));$$

$$(a_1 \text{ } \text{ }_2 (a_3 \text{ } \text{ }_1 a_2));$$

$$3) T_3 = (\text{I } (2 ( [\text{больш}] \dots \text{жИм}) \text{ I } ([\text{озер}] \dots \text{Им})) \\ [, \text{которое}] ([\text{наход}] ) \text{ в } 2 (2 ( [\text{больш}] \dots \text{жПредл}) \\ \text{I } ([\text{лес}] \dots \text{Предл}) ));$$

$$\cdot (a_1 \text{ } \text{ }_2 a_3).$$

Текст  $T_3$  образован из текста простого  $\psi$ -выражения  $(a_3 \text{ } \text{ }_1 a_2)$  при помощи текстов  $T_1$  и  $T_2$ . При этом в текст  $T_1$  в двух местах введены новые значения одного информационного признака.

4) Наконец, текст

$$T_4 = (\text{I } ([\text{Егор}] ) ([\text{двига}] ) [\text{к}] 2 (\text{I } (2 ([\text{больш}] \dots \text{жДат}) \\ \text{I } ([\text{озер}] \dots \text{Дат})) [, \text{которое}] ([\text{наход}] ) [\text{в}] \\ 2 (2 ( [\text{больш}] \dots \text{жПредл}) \text{ I } ([\text{лес}] \dots \text{Предл}))) )$$

получается по тексту  $\psi$ -выражения  $(a_1 \text{ } \text{ }_2 a_3)$ , где вместо подтекста второго аргумента стоит текст  $T_3$ , полученный из текста  $T_3$  путем изменения значения одного информационного признака в двух местах. Словоформы по тексту  $T_4$  синтезируются аналогично примеру 4.

Преимуществом первого из описанных алгоритмов является то, что (сырой) текст, синтезированный к  $i$ -ому шагу, имеет значения информационных признаков, которые в ходе дальнейшего синтеза не подлежат изменению. Изменения в процессе синтеза могут касаться только заданных сырых текстов простых реляционных выражений.

Преимуществом второго алгоритма является одновременное выполнение анализа исходного  $\psi$ -выражения и синтаксического этапа синтеза. При синтезе сложных текстов следует предпочитать первый алгоритм, где число просмотров уже синтезированного куска текста оказывается меньшим.



## Литература

1. Брябрин В.М. Диалоговая информационно-логическая система. М., изд. Научного Совета по комплексной проблеме "Кибернетика" АН СССР, 1977.
2. Виноград Т. Программа, понимающая естественный язык. М., 1976.
3. Ежкова И.В., Поспелов Д.А. Принятие решений при нечетких основаниях. I. Универсальная шкала. Изв. АН СССР. Техническая кибернетика, 1977, 6, 3-II.
4. Ежкова И.М., Поспелов Д.А. Принятие решений при нечетких основаниях. II. Схемы вывода. Изв. АН СССР. Техническая кибернетика, 1978, 2, 5-II.
5. Ефимов Е.И., Поспелов Д.А. Семиотические модели в задачах планирования для систем искусственного интеллекта. Изв. АН СССР. Техническая кибернетика, 1977, 5, 60-68.
6. Койт М. Об одном семантическом алгоритме. - Уч. зап.Тартуского гос. ун-та, 1975, 374, 229-237.
7. Литвинцева Л.В. Вопросно-ответная система, воспринимающая информацию на нормализованном естественном языке. Изв. АН СССР, Техническая кибернетика, 1977, 2, 39-46.
8. Мальковский М.Г. Программы, понимающие естественный язык. "Обработка символьной информации", 1973, 1, 73-II5.
9. Поспелов Г.С., Поспелов Д.А. Влияние методов теории искусственного интеллекта на решение традиционных задач управления. М., изд. Научного Совета по комплексной проблеме "Кибернетика" АН СССР, 1977.
10. Поспелов Д.А. Семиотические модели: успехи и перспективы. "Кибернетика", 1976, 6, II4-I23.
11. Сенин Г.В. Взаимодействие с базой данных на естественном языке. М., изд. Научного Совета по комплексной проблеме "Кибернетика" АН СССР, 1977.
12. Sandewall, E. Formal Methods in the Design of Question-Answering Systems. "Artif. Intel.", 1971, 2, 129-145.
13. Schank, R.C. Conceptual Information Processing. Fundamental Studies in Computer Science. Volume 3. North-Holland/American Elsevier, 1975.

14. Scott, A.C., Clancey, W.J., Davis, R., Shortliffe, E.H.  
Explanation Capabilities of Production-Based Consultation Systems. Stanford Heuristic Programming Project, Memo HPP-77-1, Computer Science Department, Report N° STAN-CS-77-593. Stanford University, 1977.

TWO ALGORITHMS FOR TRANSLATING FROM  $\Psi$ -LANGUAGE  
INTO RUSSIAN

M. Koit

S u m m a r y

Two algorithms have been put forward which analyze the tree of  $\Psi$ -expression [7] (top down and bottom up, respectively) and synthesize the text of the expression in Russian. The process of synthesis consists of two stages: the first forms a complex of information markers/attributes that are replaced by their lexical entries during the second stage.

# ОБ АВТОМАТИЧЕСКОМ МОРФОЛОГИЧЕСКОМ АНАЛИЗЕ ОГРАНИЧЕННОГО ЕСТЕСТВЕННОГО ЯЗЫКА

С.Р. Литвак, Т.А. Роосмаа, М.Э. Салувер, Х.Я. Ним

## Введение

В настоящей работе описывается подсистема морфологического анализа, являющаяся частью разрабатываемой вопросно-ответной системы, понимающей ограниченный эстонский язык. Морфологический анализ входного текста производится для правильного отождествления различных форм одного и того же слова и получения грамматической информации, необходимой для проведения последующего синтаксического анализа.

Морфологический анализ в разработанной подсистеме осуществляется с помощью словаря основ и ряда вспомогательных таблиц. Если слово имеет несколько различных форм основ, то в словарь включаются все формы.

## 1. Об особенностях морфологического строя эстонского эстонского языка

1.1. Система склонения. Отличительной чертой системы склонения эстонского языка является наличие 14 падежей, при этом лишь часть падежных форм образуется строго по парадигме. Падежные окончания в большинстве случаев одинаковы у всех склоняемых слов, как в единственном, так и во множественном числе. Формы множественного числа отмечаются наличием признака множественного числа перед падежным окончанием. Грамматический род отсутствует.

Главными для образования парадигмы склонения падежами являются номинатив, генетив, партитив и иллатив единственного числа и генитив, партитив и иллатив множественного.

Ниже приведен пример образования некоторых падежных форм от слова "jõgi" (река).

Падеж	Вопрос	Единств. число	Множеств. число
Номинатив	кто? что?	jõgi	jõe/d
Генитив	кого? чего	jõe	jõe/de
Партитив	кого? что?	jõe	jõe/sid
Иллатив	в кого? во что?	jõe/sse jõkke	jõe/de/sse jõe/de/s
Инессив	в ком? в чем?	jõe/s	jõe/de/s

Падежные окончания выделены. Номинатив множественного числа, а также все следующие падежи единственного числа, за исключением партитива, образуются от формы генитива единственного числа. Генитив множественного числа образуется от партитива единственного числа и служит основой для форм следующих падежей множественного числа за исключением партитива.

Кроме регулярно образованных падежных форм, употребляются еще параллельные формы: в единственном числе в иллативе, а во множественном - с иллатива по аблатив. Параллельные формы образуются во множественном числе от партитива. Пример образования множественного числа для слова "rattad" (колёса):

Падеж	Множественное число
Номинатив	rattad
Генитив	ratas/te
Партитив	rattai/d
Иллатив	ratas/te/sse, rattai/sse
Инессив	ratas/te/s, rattai/s

1.2. Система спряжения. В эстонском языке глаголы имеют четыре наклонения: изъявительное, повелительное, сослагательное и косвенное. Категория вида отсутствует. Имеется настоящее и прошедшее время, последнее имеет три различных формы - имперфект, перфект и плюсквамперфект (перфектные формы являются сложными).

Парадигма глагола включает два инфинитива и два причастия: 1. *ma*-инфинитив, имеющий пять форм (-*ma*, -*mas*, -*mast*, -*maks*, -*mata*); 2. *da*-инфинитив, имеющий две формы (-*da*, -*des*); 3. причастия настоящего и прошедшего времени, имеющие формы личного и безличного залогов:

Время	Личный залог	Безличный залог
Настоящее	-v (lugev) читающий	-tav (loetav) читаемый
Прошедшее	-nud (lugenud) читавший	-tud (loetud) (про)читанный

В эстонском языке очень распространены изменения в основе слова, происходящие и при склонении, и при спряжении. При спряжении необходимо знать основные формы глаголов, имеющие нерегулярное образование, все остальные формы образуются от основных регулярным образом. Основными формами глагола явля-

ются: 1. *ma*-инфинитив, 2. *da*-инфинитив, 3. форма первого лица единственного числа настоящего времени изъявительного наклонения, 4. причастие прошедшего времени в безличном залоге.

1.3. Ограничения на входной текст. Текст, подаваемый на вход описываемой системы автоматического морфологического анализа, в настоящее время должен удовлетворять следующим требованиям: 1. предложения должны быть в личном залоге изъявительного наклонения, 2. рассматриваются только простые слова без послелогов.

## 2. Общая схема морфологического анализа

В описываемой системе морфологического анализа используется словарь и две вспомогательные таблицы.

2.1. Структура словаря. Словарь слов представляет собой таблицу, в которой каждая строка соответствует определенной основе. В строке имеются три зоны: - морфологическая, синтаксическая и семантическая, состоящие из наборов полей. В морфологическую зону входят поля основы слова и номеров входов во вспомогательные таблицы, в синтаксическую - поля инфинитивной (начальной) формы слова и части речи. Семантическая зона состоит из единственного поля, в котором записывается семантическая категория данного слова.

Участок словаря для слов "*mārka*" (замечать) и "*naine*" (женщина) приведен ниже.

Ключ	Часть речи	Сем. категория	Номера	
<i>mārga</i>	глагол	восприятие	2,4	<i>mārka</i>
<i>mārka</i>	глагол	восприятие	1,3	<i>mārka</i>
<i>naine</i>	сущ.	лицо	1	<i>naine</i>
<i>naise</i>	сущ.	лицо	2	<i>naine</i>
<i>naist</i>	сущ.	лицо	3	<i>naine</i>
<i>naisi</i>	сущ.	лицо	4	<i>naine</i>
<i>naiste</i>	сущ.	лицо	5	<i>naine</i>

2.2. Структура вспомогательных таблиц. Вспомогательные таблицы создаются отдельно для склоняемых и спрягаемых слов. Входной ключ образуется из номера, получаемого с помощью словаря и выделенного окончания слова. С помощью вспомогательных таблиц получается вся недостающая грамматическая информация, необходимая для дальнейшего синтаксического ана-

лиза: число и падеж для склоняемых слов, число, лицо, время, инфинитив, утверждение-отрицание и наклонение для спрягаемых. Пример участка таблицы для склоняемых слов:

Ключ	Число	Падеж
2 SSE	ед.	Иллатив
5 LE	множ.	Аллатив

для спрягаемых слов

Ключ	Число	Лицо	Отр.-утвер.	Время	Наклон.
3 N	ед.	I	утв.	наст.	изъявит.
1 SIN	ед.	I	утв.	имперф.	изъявит.

### 3. Алгоритм морфологического анализа

- Шаг 1. Прочитать следующее слово.
- Шаг 2. Поиск в словаре основ. Если соответствующая основа найдена, то перейти к шагу 3, иначе - к шагу 9.
- Шаг 3. Прочитать из словаря содержимое морфологической и семантической зон.
- Шаг 4. Является ли слово спрягаемым? При положительном ответе перейти к шагу 5, в противном случае - к шагу 12.
- Шаг 5. По ключу, состоящему из номера входа и выделенному окончанию слова из вспомогательной таблицы для спрягаемых слов, прочитать необходимую информацию. Перейти к шагу 6.
- Шаг 6. Проверка на конец предложения. При положительном ответе перейти к шагу 7, иначе - к шагу 1.
- Шаг 7. Было ли в предложении отрицание? Если да, то перейти к шагу 8, иначе - к шагу 16.
- Шаг 8. Найти первый за отрицанием глагол, отметить отрицание. Перейти в шаг 16.
- Шаг 9. Проверить количество букв в отдельном окончании, если их меньше четырех, то перейти к шагу 10, иначе - к шагу 17.
- Шаг 10. Отделить следующую букву с конца.
- Шаг 11. Запомнить окончание. Перейти к шагу 2.
- Шаг 12. Является ли слово союзом или отрицательной частицей? Если да, то перейти к шагу 13, иначе - к шагу 15.
- Шаг 13. Является ли слово отрицательной частицей? При положительном ответе перейти к шагу 14, иначе - к шагу 6.
- Шаг 14. Запомнить номер слова. Перейти к шагу 6.

- Шаг I5. По окончании и номеру входа в таблицу прочитать из вспомогательной таблицы для склоняемых слов необходимую информацию. Перейти к шагу 6.
- Шаг I6. Передать полученную в ходе анализа информацию на вход синтаксического анализатора. Конец.
- Шаг I7. В тексте обнаружено незнакомое слово. Прерывание работы.

#### 4. Реализация

Вышеописанная подсистема автоматического морфологического анализа реализована на языке ПЛ/I ОС ЕС ЭВМ. Объем программ составляет 350 операторов ПЛ/I. Основным ограничением реализации в настоящее время является запрещение омонимии, вызванное экспериментальным характером системы. Словарь размещается на магнитных дисках в виде файла с индексно-последовательной организацией. Объем словаря при отладке подсистемы составлял 250 основ.

#### AUTOMATIC MORPHOLOGICAL ANALYSIS OF LIMITED NATURAL LANGUAGE

S. Litvak, T. Roosmaa, M. Saluveer, H. Õim

#### S u m m a r y

A subsystem of automatic morphological analysis (part of a natural language based question-answering system being developed at Tartu State University) is described. The morphological analysis is executed with the help of a lexicon of word stems and two auxiliary tables - one for nominals, the other for verbs. When there exist different forms of one and the same word stem all these forms have been included in the lexicon as separate units.

# ПОДСИСТЕМА АВТОМАТИЧЕСКОГО СИНТАКСИЧЕСКОГО АНАЛИЗА ДЛЯ ЭКСПЕРИМЕНТАЛЬНОЙ ВОС

С.Р. Литвак, Т.А. Роосмаа, М.Э. Салувер, Х.Я. Нийм

Подсистема автоматического синтаксического анализа является следующей ступенью ЛингП, морфологический анализатор которого был нами описан в [1].

На вход синтаксического анализатора поступает предложение, в котором для каждого слова установлена следующая синтаксически ревалентная информация: 1) часть речи; 2) для существительных и прилагательных - число и падеж; 3) для глаголов - является ли данная форма личной формой, какой-нибудь формой инфинитива или причастия; 4) каждому слову приписана одна или несколько семантических категорий.

Основную часть синтаксического анализатора составляют так называемые "правила попарной совместимости". Остановимся коротко на понятии "попарная совместимость". На основе линейного расположения слов в предложении для каждого слова составляется множество всех возможных пар (синтагм), членом которого данное слово может являться. В результате получается матрица совместимых пар. Основой составления этой матрицы служат правила синтаксического анализа. Эти правила имеют следующий вид:

Если

Слово I  
MORF:M1  
SEM:S1

и слово 2

MORF:M2  
SEM:S2

то

слово  $\xrightarrow{R}_i$  слово  $j$  ,

где R обозначает глубинно-синтаксическую связь между анализируемыми словами,  $M_i$  представляет множество требуемых морфологических признаков слова, а  $S_j$  представляет требуемую семантическую категорию (или категорий) слова. Например, если

слово I  
MORF: PS=VERB  
SEM:SUP=COMMACTION



слово 2

MORF: PS = NOUN/PRONOUN

CASE = C8

SEM1: SUP = LIVBEING,

то слово 1  $\xrightarrow{\text{REC}}$  слово 2,

где PS обозначает часть речи, SUP - семантическое надпонятие, CASE - падеж, REC - тип связи (в данном примере RECIPIENT).

Одна из отличительных черт выбранного нами варианта синтаксического анализа заключается в том, что мы стараемся "одним заходом" дойти до как можно более глубинного представления структуры предложения (до представления, которое мы называем синтактико-семантическим). С этой целью при установлении отношений между словами регулярно используются семантические категории слов (надпонятия). Эти категории являются весьма эффективным средством разрешения синтаксической эмонимии. Например, глагол "käsikima" "приказывать" может сочетаться с существительным в адессиве с падежным окончанием - 1. В зависимости от того, обозначает ли это существительное живое существо (LIVBEING), место (LOCATION) или время (TM), мы получим три разных синтагмы:

käskis tüdrukul	приказал девочке
käskis tänaval	приказал на улице
käskis hommikul	приказал утром.

Для того, чтобы различать эти три случая, нам и понадобятся семантические категории, которые приписаны словам морфологическим анализатором.

Синтаксический анализатор, установив, что какая-то пара слов допускается имеющимися правилами, выдает возможное глубинно-синтаксическое отношение между ними. Так, из приведенного примера для первой пары роль существительного будет играть REC, для второй LOC, для третьей - TM. Набором правил предусмотрены всевозможные парные комбинации слов, включенные в словарь VOC-a.

Устанавливаемые глубинно-синтаксические отношения обычно имеют четкую семантическую значимость, т.е. они могут быть перенесены в семантическое представление предложения. Но бывают и случаи, когда на данном этапе анализа не удастся определить точное содержание отношения, связывающего два сло-

ва. Это бывает прежде всего в случае связывания двух существительных: *kooli direktor* "директор школы", *küberneetika kateeder* "кафедра кибернетики" и т.д. В этих случаях для определения содержательной связи между словами нужно учитывать не только какие-нибудь общие категории слов, а уже более полные данные об объектах, обозначаемых соответствующими словами, т.е. фреймы этих слов. В таком случае на этапе синтаксического анализа между словами устанавливается отношение **ATR** которое на этапе семантического анализа подвергается содержательной интерпретации.

В качестве примера рассмотрим предложение *Programmeerimise kateedri koosolek toimus teisipäeval kateedri ruumis* (совращение кафедры программирования состоялось во вторник в помещении кафедры). Пронумеровав все слова предложения и определив всевозможные их отношения в данном предложении, мы получим информацию в виде следующей таблицы (табл. I).

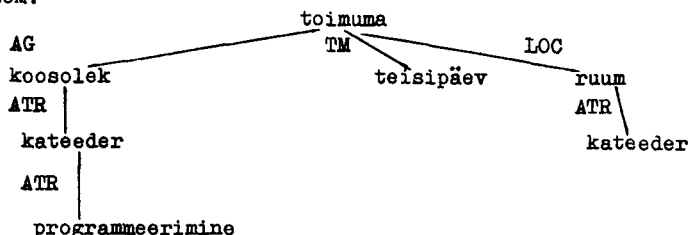
Таблица I

	I	2	3	4	5	6	7
1	X						
2	ATR	X					
3	*ATR	ATR	X				
4			AG	X	TM		LOC
5					X		
6	*ATR	*ATR				X	
7	*ATR	*ATR				ATR	X

Как видно, в таблице зафиксированы все допустимые (в глубинно-синтаксическом смысле) связи, в том числе и многие неверные (отмеченные звездочкой). Для устранения этих связей (т.е. избавления от "шума") и для получения, по возможности, однозначной структуры применяются разные фильтры/ограничения. Среди них есть такие хорошо известные требования, как требование проективности результирующей структуры (непроективные структуры в анализируемом подязыке не рассматриваются, хотя в эстонском языке, благодаря свободному порядку слов, непроективные структуры встречаются весьма часто). Есть и более "прагматические" фильтры, которые указывают, которую из воз-

возможных связей следует предпочесть. Так, в сочетании "прилагательное + существительное" в случае нескольких потенциальных "хозяев" у прилагательного в качестве хозяина выбирается ближайшее существительное среди возможных (правило близости). Это же правило применяется и при анализе сочетаний "существительное в генитиве + существительное".

Глубинно-синтаксическую структуру предложения можно представить в виде дерева зависимостей. Структуру рассмотренного предложения например, можно изобразить следующим графом.



Полученная в результате синтаксического анализа глубинно-синтаксическая структура (или структуры, если правилами допускается несколько равносильных вариантов структур предложения) является входом для следующего семантического анализа и интерпретации предложения, в результате которого строится семантическое представление предложения (экземпляр фрейма, соответствующий описанному в данном предложении событию или ситуации), формально представляющее собой семантическую сеть.

### Литература

1. С.Р. Литвак, Т.А. Роосмаа, М.Э. Салувеэр, Х.Я.Нийм. Об автоматическом морфологическом анализе ограниченного естественного языка (наст. сборник).

A SUBSYSTEM OF AUTOMATIC SYNTACTIC ANALYSIS  
FOR AN EXPERIMENTAL QUESTION-ANSWERING SYSTEM

S. Litvak, T. Roosmaa, M. Saluveer, H. Õim

S u m m a r y

Being the second one in a series of works this paper is devoted to the description of a natural language based question-answering system under development at Tartu State University.

The core of the syntactic analyzer constitute the so-called "rules of consistency in pairs". For every word of a given sentence a set of possible pairs with the other words of the sentence (syntagms) is compiled. The result of the syntactic analysis can be expressed as a matrix of consistent pairs. Rules of syntactic analysis serve as the basis for this matrix. To get rid of the "noise" in the matrix various types of filters/restrictions are used.

## НОРМАТИВНОЕ ПОВЕДЕНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

Д.А. Поспелов, В.А. Шустер

1. Введение. При создании интеллектуальных систем, способных автономно принимать целесообразные решения по организации деятельности во внешнем мире, встает задача определения норм поведения искусственной системы, допустимых при ее функционировании. Без четкого и формального задания нормативных ограничений и условий мы не сможем гарантировать того, что робот не будет пользоваться принципом: "достижение цели оправдывает любые средства ее достижения".

Три закона робототехники, сформулированные А. Азимовым /1/ и ставшие широко известными как читателям научно-фантастической литературы, так и специалистам в области систем искусственного интеллекта, во-первых, не являются формальными, а во-вторых, не охватывают сколь-нибудь полно всю область нормативного поведения\*.

Положение усугубляется еще и тем, что в психологии личности и социальной психологии отсутствуют какие-либо попытки формализации моделей поведения и точного определения понятий, используемых при словесном описании поведенческих ситуаций и действий людей в этих ситуациях.

В этой работе мы опишем аппарат, который, как нам представляется, пригоден для формализации нормативного поведения и задания необходимых законов такого поведения для роботов и других интеллектуальных систем.

Работа содержит три части.

В первой части вводится формализация понятия поступка и

---

\* Напомним эти законы:

1. Робот не может причинить вред человеку или своим бездействием допустить, чтобы человеку был причинен вред.
2. Робот должен повиноваться всем приказам, которые отдает человек, кроме тех случаев, когда эти приказы противоречат первому закону.
3. Робот должен заботиться о своей безопасности в той мере, в какой это не противоречит первому и второму законам.

**обсуждаются** некоторые вопросы, связанные с ней. Во второй части приводятся результаты психологических экспериментов, а в третьей части дается анализ возможностей введенного нами аппарата для решения проблемы формализации нормативного поведения.

2. **Формализация.** Введем ряд важных для дальнейшего понятий. Будем предполагать, что во внешнем мире действуют  $d$  активных индивидуумов  $I^1, I^2, \dots, I^d$ . Если  $i$ -ый индивидуум в некоторый момент времени совершает осознанное действие, предполагающее реакцию на это действие со стороны других  $I^j$  или внешнего мира, то мы будем говорить, что  $I^i$  совершает поступок  $\pi^i$ . Этот поступок происходит в индивидуальной для  $I^i$  поведенческой ситуации  $S_K^i$ , где  $K$  есть номер, произвольно приписываемый ситуации для удобства исследования. В  $S_K^i$  входит совокупность всех сведений о состоянии внешнего мира и других индивидуумов и весь личный поведенческий опыт, накопленный  $I^i$ .

В дальнейшем мы ограничимся лишь таким случаем, когда один индивидуум совершает одно действие, а остальные только реагируют на это действие. Поэтому верхний индекс в ситуациях и поступках мы условимся опускать.

В этой работе мы рассмотрим чрезвычайно упрощенную модель поступков. Однако многие принципиальные моменты будущей теории нормативного поведения интеллектуальных систем оказывается возможным исследовать даже на подобных простых моделях.

Для формализации воспользуемся описанием поступка  $I$  в виде двух графов, каждый из которых содержит шесть вершин, помеченных буквами  $I, H, T, N, U, S$ . Эти вершины соответственно означают: Я или лицо, совершающее действие, чей поступок оценивается исследователем; ОН или лицо, на которое непосредственно направлено действие, совершаемое  $I$  (пол  $H$  при этом не важен, кроме того  $H$  может означать и целую группу лиц); ОНИ - круг лиц, которые персонафицированы для  $I$ , это, как правило, семья, круг близких друзей, товарищей по работе и т.п.; ПРИРОДА - реальная физическая среда, в которой протекает действие  $I$ ; ЧЕЛОВЕЧЕСТВО - совокупность морально-этических норм и обычаев, принятых в данном социуме, к которому принадлежит  $I$ ; ОБЩЕСТВО - свод зафиксированных законов и постановлений, регулирующих жизнь данного социума как внутри, так и в рамках более крупного социума.

Эти шесть вершин, по-видимому, составляют тот минимум, который необходим при анализе поступков.

Вершины в графах будут соединяться дугами двух типов: положительными (сплошные стрелки) и отрицательными (пунктирные стрелки). На дугах разрешается проставлять веса, указывающие в некоторых условных единицах силу положительных и отрицательных воздействий.

Один из двух графов будет использоваться для описания планируемого действия, а второй - его реализации.

На рис. 1 показана пара таких графов. Левый граф соответствует плану, а правый - его реализации. Содержательно эти графы описывают следующее: лицо, планирующее действие, предполагает, что если будет сделано положительное действие с весом  $\varphi_1$ , то, реагируя на это, ОН ответит положительным для Я действием с весом, не меньшим  $\varphi_1$ . Кроме того, Я ожидает, что реакция персонифицированного для него окружения будет положительной с весом  $\varphi_2$ , а морально-этические нормы, принятые в данном обществе, оценивают его действие как вполне хорошее с некоторым субъективным весом  $\varphi_3$ .

Правый граф на рис. 1 показывает, что планирование было произведено плохо. После совершения действия Я неожиданно для себя получил от НЕГО отрицательное воздействие с весом  $\varphi'_1$ . Ближайшие друзья и знакомые дали также отрицательную оценку его действию с весом  $\varphi'_2$ . И лишь реакцию ЧЕЛОВЕЧЕСТВА Я предугадал правильно.

Исследователь W мог бы оценить поступок I как недалёковидный (если бы он описывал поступок H, то мог бы назвать его неблагодарным). Легко придумать такую жизненную ситуацию, в которой графы, показанные на рис. 1, действительно, реализуются. Например, I проявляет участие к H, ожидая от него благодарности за это участие, но H, считая действия I бестактными, отвечает на них грубостью. Окружение T осуждает I за то, что он "мечет бисер перед свиньями" и I остается утешаться лишь моральным удовлетворением от своей доброты.

Между словесными оценками, используемыми в человеческой практике для описания характера поступков, и графовыми описаниями введенного нами типа существует определенная корреляция. Например, для двух графов, отражающих поступок I, на рис. 1 мы уже использовали такую оценку как "недалёковидный поступок". На рис. 2 показана еще одна пара графов, которая, по-видимому, хорошо коррелирует со словесной оценкой "коры-

стный поступок". Изучение этой корреляции представляет большой интерес, т.к. описание нормативного поведения интеллектуальных систем в настоящее время возможно лишь на уровне словесных описаний, а переход к графовым представлениям позволит внести необходимые формализмы в эти описания. На языке графовых описаний хорошо решается задача типологизации поступков (в работе /3/ в связи с этим вводятся фреймы поступков). Кроме того, от этих описаний несложно перейти к секвенциальным записям. Например, графы, показанные на рис.2, эквивалентны двум секвенциям следующего вида:

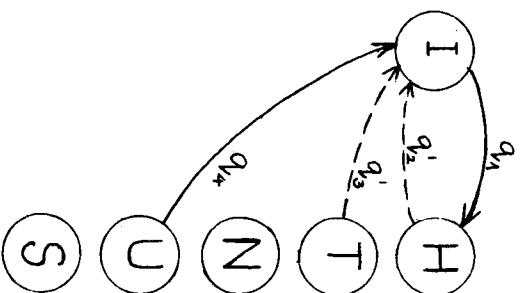
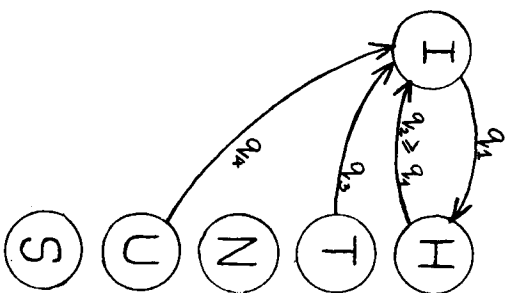
$$D_I(q_1) \Rightarrow D_H(q_2 \gg q_1) \& D_U(-q_1)$$

$$D_I(q_1) \Rightarrow D_H(q'_2 \gg q_1) \& D_N(-q_3 \ll q_1) \\ \& D_U(-q_4)$$

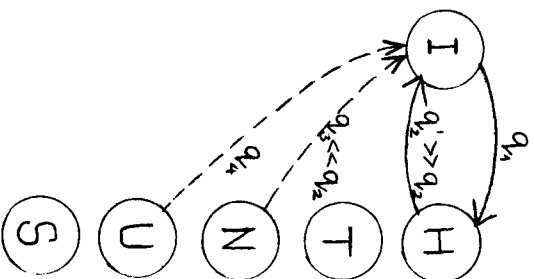
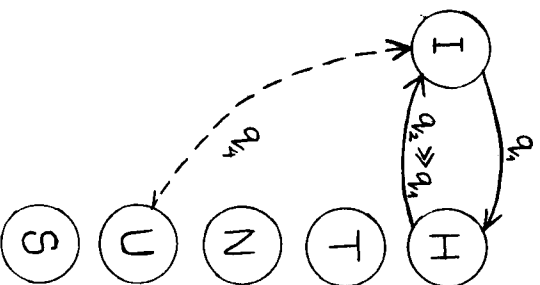
Здесь  $D$  есть предикатный операторный символ, а  $q_i$  играют роль обобщенных предикатных переменных. В рамках этой работы мы не будем затрагивать проблемы, связанные с построением специальной алгебры графов поступков или логики секвенциальных записей. Не будем мы рассматривать и описания, использующие в секвенциях модальные операторы типа: возможно, необходимо, полагается и т.п., которые более адекватны реальным поступкам. Нас в последующих разделах будет интересовать фундаментальный для жизненности рассматриваемого подхода к формализации нормативного поведения интеллектуальных систем (не человека!) вопрос о корреляции графовых формализмов и словесных оценок поступков.

3. Психологические эксперименты. До начала экспериментов была проделана следующая подготовительная работа. Из словаря русского языка были выписаны все слова, используемые для словесной оценки поступков и для каждого слова (их число превысило 200). Коллектив экспериментаторов построил графовое описание поступка, который, по мнению экспериментаторов, способствует данной словесной оценке. При этом делались попытки выделить в графах минимальное число связей, чье присутствие было необходимо для корреляции со словесной оценкой. Другими словами, осуществлялась попытка построения определя-





P n c. I.



P n c. 2.

ющих фреймов для словесных оценок. Например, для оценки "корыстный" бралась не пара графов, приведенных на рис. 2, а такая пара, в которой в первом графе нет связи, идущей от вершины  $\mathcal{N}$  к вершине I.

По полученным фреймам строились словесные описания ситуаций и действий. Примеры этих описаний для одной группы экспериментов приведены в приложении I. Эти описания использовались двояко. В экспериментах первого типа испытуемым предлагалось для каждого описания дать словесную оценку поступка, выраженную одним словом. Предлагалось указывать синонимичные оценки и близкие по смыслу слова, если, по мнению испытуемого, они возможны. Допускались произвольные комментарии к этим оценкам.

В экспериментах второго типа испытуемым вместе с описаниями давался список слов, набранный из оценок, сделанных экспериментаторами. Цель испытуемого состояла в установлении поэлементного соответствия данных ему описаний со словами списков. Как и в первом случае испытуемый мог писать любые комментарии к своим решениям (например, считать, что в списке оценок нет подходящего слова для оценки данного поступка, или предлагать свои слова - оценки, отсутствующие в исходном списке).

Эксперименты проводились, в основном, со студентами математиками, студентами инженерного и гуманитарного профиля, некоторыми группами научных работников и школьниками. Всего в экспериментах обоих типов участвовало 125 человек (в первом - 69 человек, во втором - 56 человек).

Результаты I-го эксперимента показали, что среди первых 4-х наиболее часто названных слов всегда были слова, положенные в основу описания поступка (см. таблицу I).

В основу списков для экспериментов второго типа были положены предыдущие результаты, т.е. мы взяли наиболее предпочитаемые слова и добавили к ним слова, достаточно близкие по смыслу. Таким образом мы получили список из 10 слов, которые были расположены в случайном порядке (см. приложение 2). Испытуемым были предложены тексты № 6-15, по которым было меньше всего разногласий. Результаты, приведенные в таблице 2, показывают, что из наиболее предпочитаемых слов по текстам № 6, 7, 11, 12, 15 были названы те четыре слова, которые называли испытуемые в эксперименте I типа. По текстам № 8, 9, 10, 13, 14, из 4-х основных слов, названных в предыдущем

эксперименте, были отмечены 3 слова. Кроме того, испытуемых просили отметить доминантное слово. Результаты, приведенные в таблице 2, показывают, что, кроме текстов № 9, 12, во всех остальных случаях мы получили одно из тех слов, которые были положены в основу эксперимента.

На данном этапе мы пока не будем утверждать, что именно фреймы помогли испытуемым опознать поступки. Но по крайней мере можно утверждать, что 1) фреймы позволяют дать такое описание, которое большинство испытуемых интерпретирует однозначно; 2) мы получили достаточно точные словесные описания поступков и именно это доказывает результаты экспериментов.

4. Заключительные замечания. Хорошая корреляция между словесными оценками поступков, традиционно используемыми людьми, и предложенными графовыми описаниями поступков позволяет считать, что удалось создать приемлемую для интеллектуальных систем совокупность формальных понятий, в рамках которых можно описывать и задавать нормативное поведение.

В ближайшем будущем необходимо решить проблему "сцепления" поступков в сценарий поведения, научиться оценивать не отдельные поведенческие акты, а последовательности таких актов, связанных мотивационными связями. Теория сценариев поведения пока еще находится в зачаточном состоянии, хотя некоторые результаты, полученные здесь, уже могут принести пользу /4/. На их основе можно будет строить теорию нормативных поведенческих сценариев, используя результаты по нормативному описанию поступков. Этот путь нам представляется весьма перспективным.

Т а б л и ц а    I

Результаты эксперимента первого типа

№ текста	Слова, положенные в основу описания	Полученные результаты	
I :	2	:	3
I. <u>Смелый</u> <sup>I</sup>			
	храбрый,	Смелый	- 32 чел. из 69;
	отважный	волевой	- 23 чел. -"
		мужественный	- 22 чел. -"
		целеустремленный	- 20 чел. -"
		храбрый	- 10 чел. -"

<sup>I</sup> Примечание: - подчеркиваются слова, которые для данного текста являются доминантными.

Продолжение табл. I

I :	2	:	3
2. <u>Великодушный</u>			
благородный	добрый	- 27	чел. из 69
добрый	великодушный	- 25	чел. -"
	бескорыстный	- 19	чел. -"
	благородный	- 16	чел. -"
3. <u>Недальновидный</u>	Безрассудный	- 11	чел. из 69
безрассудный	необдуманный	- 9	чел. -"
	попешный	- 2	чел. -"
	опрометчивый	- 2	чел. -"
4. <u>Мужественный</u>	Мужественный	- 31	чел. из 69
благородный	смелый	- 25	чел. -"
	геройский	- 19	чел. -"
	благородный	- 11	чел. -"
5. <u>Отчаянный</u>	Безрассудный	- 14	чел.
безрассудный	отчаянный	- 9	чел.
	настойчивый	- 8	чел.
6. <u>Геройский</u>	Геройский	- 33	чел. из 69
самоотверженный	мужественный	- 20	чел. -"
подвиг	самоотверженный	- 16	чел. -"
	подвиг	- 9	чел. -"
7. <u>Корыстный</u>	Подхалимский	- 19	чел. из 69
хитрый	корыстный	- 17	чел. -"
	расчетливый	- 9	чел. -"
	лицемерный	- 7	чел. -"
8. <u>Коварный</u>	Подлый	- 38	из 69
неискренний	коварный	- 4	-"
неизменный (под-	злойный	- 8	-"
лый), лживый	хитрый	- 4	-"
9. <u>Несправедливый</u>	Подлый	- 42	из 69
пристратный	злойный	- 10	-"
	клеветнический	- 5	-"
	нечестный	- 3	-"
10. Бестактный	Нетактичный	- 22	из 69
(нетактичный),	необдуманный	- 12	-"
неловкий	неловкий	- 7	-"
	невнимательный	- 5	-"
II. <u>Бескорыстный</u>	Бескорыстный	- 41	из 69
великодушный,	доброжелательный	- 37	-"
доброжелательный	(добрый)		
	великодушный	- 10	-"
	благородный	- 5	-"

I	:	2	:	3
I2.	<u>Безжалостный</u> , бессердечный, бездущный	Равнодушный эгоистичный бессердечный (бездущный) черствый	- 14 из 69 - 11 "- - 9 "- - 10 "-	
I3.	<u>Жестокий</u>	Жестокий злойный садистский бесчеловечный	- 31 из 69 - 15 "- - 6 "- - 3 "-	
I4.	<u>Неблагодарный</u> <u>низменный</u> недостойный	Эгоистичный неблагодарный непорядочный равнодушный	- 19 из 69 - 19 "- - 6 "- - 5 "-	
I5.	<u>Принципиальный</u> решительный смелый	Принципиальный честный справедливый смелый решительный	- 28 из 69 - 25 "- - 20 "- - 12 "- - 6 "-	

Мы предполагаем, что поступки  $\Pi^i$  оцениваются некоторым внешним наблюдателем - исследователем  $W$ . С точки зрения  $W$  поступок  $\Pi^i$ , определяемый для  $I^i$  ситуаций  $C_K^i$ , может происходить совсем в иной ситуации  $C_K^W$ , т.к. личный опыт  $W$  и его знания о внешнем мире могут отличаться от знаний, имеющих в распоряжении  $I^i$ . В качестве  $C_K^W$  может выступать, например, объективная поведенческая ситуация  $C_K$ , получаемая как  $C_K^i$ . Другими словами,  $C_K$  есть совокупность тех сведений, которые являются общими для всех индивидуумов, действующих одновременно во внешнем мире. В дальнейшем мы будем предполагать, что  $C_K^W$  совпадает с  $C_K$  (тем самым мы исключаем оценки поступков  $I^i$  в тех случаях, когда  $W$  знает больше, чем  $I^i$  о ситуации, в которой совершается поступок).

Как следует из нашего определения поступка, он является мотивированным действием. Более подробно о психологических аспектах, связанных с понятием поступка, мы говорить не будем. Это было сделано ранее в работе /2/. Отметим только, что всякий поступок  $\Pi_i$  состоит из мотивированного плана будущего действия, в котором прогнозируются и ответные реакции на действие, и реализации этого действия. Кроме того любой индивидуум всегда может встретиться с ситуацией, когда истин-

Т а б л и ц а 2

## Результаты эксперимента второго типа

№ текста	Слова, выбранные из списка		Доминантное слово		Примечания
1	:	2	:	3	4
6.	Подвиг	- 49 из 56	<u>Подвиг</u>	- 39 из 56	
	геройский	- 34 -"			
	самоотверженный	- 30 из 56			
	мужественный	- 25 -"			
7.	Расчетливый	- 40 из 56	<u>Корыстный</u>	- 15 из 56	Остальные слова приводятся для сравнения
	подхалимский	- 40 -"	<u>расчетливый</u>	- 13 из 56	
	лицемерный	- 24 -"	подхалимский	- 14 -"	
	корыстный	- 22 -"	лицемерный	- 5 -"	
8.	Подлый	- 41 из 56	<u>Подлый</u>	- 22 из 56	
	(низменный)		(низменный)		
	коварный	- 35 -"			
	хитрый	- 16 -"			
	нечестный	- 12 -"			
9.	Клеветнический	- 51 из 56	<u>Клеветнический</u>	- 30 из 56	
	подлый	- 35 -"			
	нечестный	- 24 -"			
	расчетливый	- 8 -"			
10.	Необдуманный	- 32 из 56	<u>Бестактный</u>	- 25 из 56	
	бестактный	- 29 -"			
	невнимательный	- 27 -"			
	непреднамеренный	- 27 -"			

I	:	2	:	3	:	4
II.	Бескорыстный	- 45 из 56		<u>Бескорыстный</u>	- 36 из 56	
	добрый	- 31 -"				
	великодушный	- 27 -"				
	благородный	- 29 -"				
I2.	Равнодушный	- 32 из 56		<u>Эгоистичный</u>	- 14 из 56	
	черствый	- 31 -"				
	бессердечный	- 29 -"				
	эгоистичный	- 47 -"				
I3.	Жестокый	- 37 из 56		<u>Жестокый</u>	- 23 из 56	
	бесчеловечный	- 36 -"				
	садистский	- 16 -"				
	злосный	- 14 -"				
I4.	Неблагодарный	- 40 из 56		<u>Неблагодарный</u>	- 22 из 56	
	непорядочный	- 29 -"				
	эгоистичный	- 25 -"				
	несправедливый	- 8 -"				
I5.	Принципиальный	- 31 из 56		<u>Принципиальный</u>	- 20 из 56	
	честный	- 29 -"				
	справедливый	- 29 -"				
	решительный	- 16 -"				

ная реакция на его действие будет отличной от прогнозируемой.

Для оценки поступка исследователь должен знать  $V^i (C_k^i)$  и объективные реакции остальных лиц, участвующих в ситуации на действие, совершенное  $i$ -м индивидуумом. Таким образом, нас будет в дальнейшем интересовать внешняя оценка действий индивидуума в определенной ситуации. Исследователь при этом стоит вне ситуации и действий индивидуумов.

### Литература

1. Азимов А. Путь марсиан. М., "Мир", 1966.
2. Захаров В.Н., Палилова С.Ф., Поспелов Д.А., Шустер В.А. О формализации структуры поведения. - Вопросы бионики. Харьков, 1979.
3. Поспелов Д.А., Шустер В.А. Фреймы, поступки и личностные характеристики. - Мат-лы конференции "Психология человеческого учения и развития". Прага, 1979.
4. Шенк Р., Абельсон Р. Сценарии, планы и знания. - Труды IV международной объединенной конференции по искусственному интеллекту, вып. 6. М., 1975.



Т Е К С Т Ы

- I. Ситуация не в пользу А и даже имеется угроза его жизни, но преодолев всякие опасения, он совершает задуманные действия.
2. А, несмотря на то, что В причинил ему в свое время неприятности, оказывает ему помощь, не ожидая, что В оплатит ему тем же.
3. А, желая помочь В, совершает такие действия, которые не оказывают помощи В, но при этом сам А имеет большие неприятности.
4. Находясь в тяжелом положении, А, рискуя своей жизнью, спасает жизнь В.
5. А, видя, что не может достичь успеха в отношениях с В, тем не менее предпринимает действия, которые могут принести самому А почти смертельный исход.
6. Рискую жизнью, А спасает жизнь людям, но при этом сам погибает.
7. Желая получить от В некоторую услугу, А оказывает ему самые разные знаки внимания.
8. Под видом доброжелательства и внимания А совершил по отношению к В такие действия, которые принесли ущерб В (но В об этом не знает).
9. А, зная, что В не виноват, тем не менее обвиняет и осуждает его.
10. Своими действиями (непреднамеренными) А ставит В в неловкое положение.
- II. А оказывает постоянно помощь В, не ожидая взамен ничего.

- I2. А, имея возможность оказать поддержку и помощь В, который в этом очень нуждается, тем не менее отказывается это сделать.
- I3. А, понимая, что причиняет значительный вред В (имеется угроза для жизни В), тем не менее продолжает это делать.
- I4. В неоднократно оказывал помощь А, но когда В понадобилась помощь, А ее не оказал.
- I5. А выступил на собрании и, невзирая ни на что, справедливо критиковал недостатки В.

ТЕКСТ 6

1. ОТЧАЯННЫЙ.
2. ГЕРОЙСКИЙ.
3. БЛАГОРОДНЫЙ.
4. САМООТВЕРЖЕННЫЙ.
5. СЕРДЕЧНЫЙ.
6. МУЖЕСТВЕННЫЙ.
7. ВЕЛИКОДУШНЫЙ.
8. ЧЕЛОВЕКОЛЮБИВЫЙ.
9. ПОДВИГ.
10. РИСКОВАННЫЙ.

ТЕКСТ 8

1. ХИТРЫЙ.
2. ЗЛОВНЫЙ.
3. НЕЧЕСТНЫЙ.
4. ЖЕСТОКИЙ.
5. МАЛОДУШНЫЙ.
6. КОВАРНЫЙ.
7. ПОДЛЫЙ.
8. ОБМАН.
9. НЕДОСТОЙНЫЙ.
10. НЕИСКРЕННИЙ.

ТЕКСТ 10

1. ЭГОИСТИЧНЫЙ.
2. БЕСЦЕРЕМОННЫЙ.
3. БЕСТАКТНЫЙ.
4. НЕПРЕДНАМЕРЕННЫЙ.
5. НЕОБДУМАНЫЙ.
6. БЕЗЗЛОВНЫЙ.
7. НЕВНИМАТЕЛЬНЫЙ.
8. ХИТРЫЙ.
9. НЕУКЛЮЖИЙ.
10. НЕРАЗУМНЫЙ.

ТЕКСТ 7

1. РАСЧЕТЛИВЫЙ.
2. ТРУСЛИВЫЙ.
3. ХИТРЫЙ.
4. КОРЫСТНЫЙ.
5. ПОДХАЛИМСКИЙ.
6. СЕВЯЛЮБИВЫЙ.
7. ХАНЖЕСКИЙ.
8. ЛИЦЕМЕРНЫЙ.
9. ОБМАН.
10. ЛЖИВЫЙ.

ТЕКСТ 9

1. ЛИЦЕМЕРНЫЙ.
2. ТРУСЛИВЫЙ.
3. КАРЬЕРИСТСКИЙ.
4. ХИТРЫЙ.
5. ЗЛОВНЫЙ.
6. ПОДЛЫЙ.
7. КЛЕВЕТИЧЕСКИЙ.
8. ЖЕСТОКИЙ.
9. РАСЧЕТЛИВЫЙ.
10. НЕЧЕСТНЫЙ.

ТЕКСТ 11

1. ЩЕДРЫЙ.
2. ВЕЛИКОДУШНЫЙ.
3. СОЧУВСТВЕННЫЙ.
4. ДОБРЫЙ.
5. УЧИВЫЙ.
6. БЕСКОРЫСТНЫЙ.
7. БЕСХИТРОСТНЫЙ.
8. БЛАГОРОДНЫЙ.
9. ПОРЯДОЧНЫЙ.
10. ГУМАННЫЙ.

ТЕКСТ 12

1. РАСЧЕТЛИВЫЙ
2. МАЛОДУШНЫЙ
3. БЕССЕРДЕЧНЫЙ
4. РАВНОДУШНЫЙ
5. ЭГОИСТИЧНЫЙ
6. КОРЫСТНЫЙ
7. ЖЕСТОКИЙ
8. ЧЕРСТВЫЙ
9. ГЛУПЫЙ
10. НЕВНИМАТЕЛЬНЫЙ

ТЕКСТ 14

1. НЕСПРАВЕДЛИВЫЙ
2. ЧЕРСТВЫЙ
3. НЕБЛАГОРОДНЫЙ
4. НЕВНИМАТЕЛЬНЫЙ
5. ЭГОИСТИЧНЫЙ
6. ХИТРЫЙ
7. ТРУСЛИВЫЙ
8. НЕПОРЯДОЧНЫЙ
9. НЕДАЛЬНОВИДНЫЙ
10. РАВНОДУШНЫЙ

ТЕКСТ 13

1. ЧЕРСТВЫЙ
2. ЭГОИСТИЧНЫЙ
3. ЦИНИЧНЫЙ
4. ЖЕСТОКИЙ
5. КОРЫСТНЫЙ
6. БЕСЧЕЛОВЕЧНЫЙ
7. КОВАРНЫЙ
8. ЗЛОБНЫЙ
9. САДИСТСКИЙ
10. ИЗДЕВАТЕЛЬСТВО.

ТЕКСТ 15

1. СПРАВЕДЛИВЫЙ
2. БЛАГОРОДНЫЙ
3. ПРИНЦИПИАЛЬНЫЙ
4. БЕЗРАССУДНЫЙ
5. РЕШИТЕЛЬНЫЙ
6. ЧЕСТНЫЙ
7. БЕСПОЩАДНЫЙ
8. СМЕЛЫЙ
9. НЕТЕРПИМЫЙ
10. ОТВАЖНЫЙ

## NORMATIVE BEHAVIOR OF INTELLECTUAL SYSTEMS

D. Pospelov, V. Shuster

### S u m m a r y

When developing intellectual systems capable of independently adopting purposeful decisions for organizing their activities in the outer world, the problem of determining the normative behavior of these systems arises. The authors say that without an exact and formal definition of the normative conditions and restrictions it cannot be guaranteed that a robot will not sometimes be tempted to proceed from the principle "The end justifies the means".

A mechanism that, according to the authors, should be suitable for formalizing normative behavior and presenting the rules that guarantee such behavior of robots and other intellectual systems is described. The first part of the paper presents the formalization of the concept of deed/act and a discussion of some problems connected with that concept. Part Two discusses the results of some psychological experiments. Part Three is devoted to the analysis of the possibilities of the given mechanism from the point of view of solving the problem of formalizing normative behavior.

## ПРОБЛЕМЫ ПОСТРОЕНИЯ АВТОМАТИЗИРОВАННЫХ ИМИТАТОРОВ ПРОЦЕССА ТВОРЧЕСТВА

В.М. Сафонов

### Введение

В современном мире проблема чтения и построения текстов с зафиксированными в них результатами творческой деятельности имеет огромное значение. Именно этим объясняется появление и существование автоматизированных систем общения на естественном языке (ЕЯ) или сокращенно, автоматизированных языковых систем (АЯС). Стремление ускорить процесс поиска привело лингвистов к использованию в АЯС одной из технических новинок века – электронных вычислительных машин (ЭВМ). После того, как несколько десятков лет тому назад были созданы ЭВМ для решения задач по их прямому назначению (вычислительные), было замечено, что ЭВМ могут решать и логические задачи. Из того обстоятельства, что в своей деятельности человек может использовать логические построения, где-то и когда-то (сейчас это установить уже не так просто) был сделан вывод о том, что проблема создания АЯС может быть сведена к проблеме обработки информации по правилам формальной логики и, следовательно, решена в рамках и методами кибернетики. Немало способствовало укреплению такого взгляда на задачи и принципы построения АЯС и то, что кибернетика уже на самых ранних этапах ее возникновения взялась решить путем логической обработки информации с учетом все возрастающего быстродействия ЭВМ и объема их памяти проблему создания искусственного разума.

Прошли годы, быстродействие и память ЭВМ достигла почти астрономических цифр, продолжая все время расти, разработаны сложные математические аппараты дискретно-логической обработки информации, но до сих пор не решена не только вышеуказанная проблема кибернетики, но и ее частная проблема – создание высокоэффективных АЯС на базе ЭВМ, или как их сейчас иногда стали называть – компьютеров. В отличие от тех кибернетических систем, которые имеют дело с неживой природой, АЯС работает непосредственно на человека, на целые коллективы в

человеческом обществе. Здесь самым существенным критерием эффективности системы является оценка ее деятельности со стороны человека.

Заметим, что все последующее изложение будет относиться к тем АЯС, в которых отсутствует четко формализованная структура построения текстов (статистические, экономические и т.д.).

Иными словами, речь будет идти о тех текстах, которые относятся к сфере творческой деятельности. Все сказанное далее будет относиться также и к весьма важному требованию, предъявляемому к современным и перспективным АЯС - творческий анализ сообщений и возможность принятия решения.

Степень многогранности и сложности проблемы создания современных АЯС можно проиллюстрировать следующим примером. Пусть необходимо при помощи АЯС прочитать рукописный текст на английском языке и на основе анализа его содержания принять и сообщить человеку некоторое решение.

Здесь необходимо решить целый комплекс проблем: распознавания образов, машинного перевода с одного языка на другой, построения умозаключений, общения с человеком на естественном языке. Для решения задач такого класса на ЭВМ надо уметь имитировать творческие процессы, т.е. необходимо создать автоматизированный имитатор процессов творчества (АИПТ).

Хотя в 50-е годы нашего столетия основоположники кибернетики и декларировали быстрое и успешное решение этих и многих других задач в рамках проблемы искусственного интеллекта, они не решены до настоящего времени.

Цель настоящей работы состоит во-первых, в том, чтобы сформулировать причины возникших трудностей и, во-вторых, обратить внимание специалистов по АЯС на метод анализа степени этих трудностей и на существование уже в настоящее время путей преодоления некоторых из них.

### О существенных особенностях мышления человека

Известно мнение, что при создании АИПТ надо решить три основные проблемы: информационного (искусственного) языка (ИЯ), носителя информации и выбора технических средств (например ЭВМ).

Такое мнение применительно к АЯС базируется на постулате

о том, что передача творческих мыслей, изложенных в тексте, от одного человека к другому возможна на уровне и средствами дискретной формальной логики. Если принять такой постулат верным, то далее оказывается, что все проблемы ИЯ сводятся к созданию "хороших" переложений (перевода) естественного языка текста на формализованный язык, обработке такого ИЯ посредством логических построений и восстановлению текста по ИЯ.

Отсюда (из такой постановки вопроса) совершенно естественно родился повышенный интерес к знаковым системам (семиотика) и стремление лингвистов заполучить в свои руки ЭВМ с возможно большим быстродействием и памятью.

Между тем, представляет огромный практический интерес взглянуть с точки зрения современного естествознания на текст, как на зафиксированную мысль человека, которая сначала была получена, а затем и изучена другим человеком, после работы мышления и в процессе творчества. Сейчас накопилось много сведений, полученных в самых различных областях знаний (неврофизиологии, психологии, физике, кибернетика), которые позволяют детализировать процесс творчества человека, процесс мышления при создании нового.

Адекватный, релевантный процесс чтения и построения текстов не может не учитывать специфику мышления человека - создателя этих текстов. Если в процессе творчества они создаются, а затем изучаются на уровне формальной логики и дискретными средствами, то при совершенствовании современных АИПТ все дело состоит, действительно, в совершенствовании ИЯ, программ обработки и компьютеров. Если такой уровень мышления человека не единственный и, может быть, даже не главный, тогда принципы построения, заложенные в современные АИПТ, являются более общих принципов. Тогда проблема ИЯ и компьютеров оказывается вложенной в проблему понимания тех уровней и тех средств человеческого мышления, средств человеческого мышления, которые участвуют в процессе создания нового и передаче его другому человеку.

Не представляет сомнений, что процессы чтения текстов, а тем более, процессы принятия решений человеком, которые ведутся на естественном языке и естественным образом, представляют творческий акт, в котором участвует мышление человека со всеми его законами, присущими процессу творчества.

Отсюда следует, что, хотим мы этого или нет, проблема по-



вышения эффективности АИПТ есть проблема понимания сущности творчества в процессе мышления человека. Задача о полноте, релевантности, точности и т.д. АЯС есть также задача о том, как человек делает нетривиальные выводы, о том, как он мыслит и принимает решения.

Накопилось много сведений о том, как происходит решение нетривиальных задач в процессе творчества у человека и какие уровни мышления и средства обработки информации в нем участвуют. Интерес к этому направлению достаточно рельефно сформулирован в [1:15]. "Примечательно, что в последнее время - в годы невиданного ранее распространения машинных систем сбора и обработки информации - человек как носитель информации и творческих идей - стал объектом многочисленных экспериментов и исследований".

Результаты таких исследований, например, Ж.Адамар [1:106] излагает следующим образом: "... говоря строго, практически не существует чисто логических открытий вмешательство бессознательного необходимо по крайней мере для того, чтобы стать отправным пунктом логической работы". Интересно высказывание А. Эйнштейна, приведенное в [1:60]. "Психическими элементами мышления являются более или менее ясные знаки или образы, которые могут быть "по желанию" воспроизведены и скомбинированы. Существует, естественно, некоторая связь между этими элементами и рассматриваемыми логическими концепциями".

Анализируя вопрос о том, присущ ли машине разум, Д.А. Шрейдер [13] отмечает; что "... ученые часто получают первоклассные результаты как раз тогда, когда размышляют о предмете без конкретной цели".

Приведенные выше высказывания специалистов в области математики, физики, философии (список мог быть и продолжен) объединяет мысль о том, что процесс творческого акта не сводится только к формально-логическим построениям в процессе обработки информации. Эта существенная особенность мышления установлена не только на уровне мнений ученых, но и современным естествознанием на уровне исследований в области нейрофизиологии.

В своей работе "Творческая активность сознания" М.Н. Морозов [9] прямо указывает следующее. "При самом общем подходе к классификации информационных процессов на уровне мозга (психики) представляется вполне обоснованным исходить из двух качественно различных форм ее протекания - вербализо-

ванной (языковой) и невербализованной (неязыковой). Такое деление отражает два основных уровня человеческой психики - субъективно осознаваемый и неосознаваемый [3, 6] и согласуется с учением И.П. Павлова о двух сигнальных системах". (стр. 122). В этой же работе М.Н. Морозов приходит к выводу, представляющему весьма существенный интерес для понимания проблемы АИПТ. "Продуктивное мышление отнюдь не может быть сведено к языковым операциям, к выбору слов и их сочетанию в правдоподобные ассоциации" и далее "Поскольку какой-то объем информации перерабатывается средствами "физиологического кода", минуя уровень языкового анализа и синтеза, можно допустить, что мышление как интегративная деятельность мозга должно включать в себя ряд соподчиненных уровней информационных процессов, каждый из которых характеризуется своим сигнальным кодом:

1) фундаментальные физиологические принципы (программы) переработки информации, которые существенно определяют психологическую структуры мысли - правила преобразования ее элементов, форму и модальность мысли и пр;

2) первичная организация текущей информации на уровне корковых интегральных центров;

3) высший корковый анализ и синтез, сооставление информации с прошлым опытом, установками и целями субъекта;

4) вербально-ассоциативная деятельность, на уровне которой осуществляется перевод (по правилам грамматики и логики) информации физиологического уровня в языковую, социально-коммуникативную форму". (стр. 129 и стр. 130).

Из вышеизложенного следует, что:

- естественный (не автоматизированный) процесс языковой коммуникации, будучи творческим процессом, не сводится только к проблеме языка и его совершенствования;

- процесс анализа сообщения и принятия решения человеком производится многоуровневым способом обработки информации, причем, формально-логическая обработка соответствует только одному и самому нижнему уровню в иерархии процесса мышления;

- моделирование и имитация творческого акта языковой коммуникации и принятия решений (прогнозирование), начатое в традиционных (известных) АИПТ только на формально-логическом уровне, может совершенствоваться только на основе все большего охвата верхних уровней иерархии (учет прошлого опыта, самообучение, ассоциативное и образное мышление, эвристика и т.д.).

Изучение процесса обработки информации во время творчества дает возможность обратить внимание еще на одну его важную особенность. В статье "Три стороны интеллекта" американский психолог Дж. Гилфорд [II: 435-437] создает объемную модель структуры интеллекта в трехмерном пространстве, по одному измерению которого располагаются различные виды операций мышления, по другому - различные виды результатов мышления, по третьему - различные виды содержания мышления. Отсюда возникает мысль о том, что "если логические процессы в машинах разворачиваются как бы в одной плоскости, то логические процессы в мозгу человека - в двух или даже в трех плоскостях" [2:96]. Заметим, что если всего несколько лет назад такая мысль не могла задержать на себе внимания из-за недостатка технических средств, то в настоящее время, в связи с новыми достижениями физики, рассмотренными в следующем разделе, она может оказаться практически полезной.

Завершая краткий обзор существенных особенностей языковой коммуникации как творческого акта, следует обратить внимание на то, что мышление человека представляет тесное единство прерывного (дискретного) и непрерывного. Сначала рассмотрим сущность этой особенности, затем ее практическую значимость и осуществимость для проблемы совершенствования АИПТ. "Обсуждая проблему врожденных ("встроенных") управления, Ф. Розенблатт пишет: "Как указал фон Нейман, кажется весьма вероятным, что мозг в своей деятельности широко использует как дискретный, так и аналоговый принципы, и, по-видимому, устройства обоих типов предопределены генетически" [9:140]. Идея о том, что деятельность мозга по обработке информации и прерывна и непрерывна, высказана академиком А.Н. Колмогоровым в одном из выступлений [2:98]. Физиологическая интерпретация этой мысли "допускает протекание на уровне нейрона сложных интегральных процессов, подобных процессу принятия решения". [9:130].

Значение этой особенности процесса обработки информации для АИПТ состоит в том, что "при воспроизведении логических процессов с помощью одной из сторон (прерывность или непрерывность) теряется нечто, присущее только их единству. Единство этих сторон в мышлении человека приводит к замедлению процесса обработки информации. Однако, это замедление компенсируется тем, что информация обрабатывается не в одном, а во многих планах, как бы в нескольких измерениях. Это каче-

ственное отличие логических процессов в мозгу человека от логических процессов в машинах находит свое выражение в ассоциативном характере мышления человека и, вероятно, лежит в основе творческого мышления". [2:98-99].

Реализация этого принципа, по-видимому, может оказаться доступной методами физической обработки информации, например, в ассоциатроне, голографе и другими принципами, которые будут рассмотрены в следующем разделе.

Таким образом, проблеме АИПТ, по-видимому, необходимо рассмотреть как часть общей проблемы моделирования мышления на более высоких уровнях, чем дискретный и формально-логический.

### О предельных возможностях АИПТ

Итак, анализ существенных особенностей мышления человека приводит к выводам о том, что в процессе творчества:

- участвуют два уровня мышления (вербализованный и невербализованный) и одновременно в разных пропорциях несколько способов обработки информации,
- обработка ведется на многомерном "носителе" информации,
- имеет место диалектическое единство прерывистого и непрерывного (дизъюнктивного и недизъюнктивного, в терминологии известного психолога А.В. Брушлинского [5]).

С учетом этих особенностей можно построить модель мышления в виде диаграммы, по горизонтальной оси которой отклонены способы обработки информации, а по вертикальной - вербализованный рассудочный уровень мышления (сознание) и выше его - невербализованный уровень разума (подсознание). На данной диаграмме формально-логический способ обработки информации располагается целиком в области нижнего (вербализованного) уровня, эвристический способ поднимается частично в верхний (невербализованный) уровень.

Все больше и больше заходят в верхний уровень на этой диаграмме последовательно способы обработки информации: использование прошлого опыта, затем - самообучение, еще выше - ассоциативный и почти целиком на невербализованном уровне - образное мышление и эмоциональное восприятие.

Поставим следующий вопрос: при непрерывном совершенствовании АИПТ, беспредельно ли будет улучшаться качество имитации или существует предел возможностей такой имитации?

Имея ввиду предложенную выше модель мышления человека, на этот вопрос можно ответить следующим образом. На начальном этапе совершенствования АИПТ при использовании того или иного способа обработки информации сначала будет происходить рост качества имитации до тех пор, пока оно не окажется равным той в доле, которую этот, например, формально-логический способ обработки информации занимает в общем объеме способов, участвующих в решении творческой задачи. При подключении следующего, например, эвристического способа этот порог качества имитации повысится, но по-прежнему будет равен той доле, которую он занимает в общем объеме обработки информации при решении этой задачи. Следует подчеркнуть, что при использовании автоматизированных имитаторов, основанных на вербализации процесса, имитируются только те части перечисленных выше способов обработки информации, которые располагаются в пределах вербализованного уровня человеческого мышления. Отсюда следует, что имитация вербализованными средствами следующих способов обработки информации (использование прошлого опыта, самообучение ассоциативно-образное мышление), далеко уходящих в невербализованный уровень будет давать все меньший вклад в верхний предел (порог) эффективности АИПТ при их непрерывном совершенствовании.

Таким образом, существует предел возможностей автоматизированных систем при имитации процесса творчества. Даже при одновременном подключении всех перечисленных выше способов обработки информации этот предел не превосходит той доли, которую занимает вербализованный уровень мышления человека в общем объеме процессов, принимающих участие в конкретном виде творчества.

Проиллюстрируем сформулированный выше принцип предельных возможностей АИПТ на примере, изложенном во введении. При автоматизированной обработке, например, машинного текста, содержащего статистическую информацию, основным способом обработки является, конечно, формально-логический. Применение ЭВМ, поэтому, может позволить не только реализовать процесс его обработки, но и превзойти человека по быстродействию. При автоматизированной обработке того же, но рукописного текста, необходимо привлечение ассоциативно-образного мышления, который может быть реализован на ЭВМ в силу сформулированного выше принципа лишь частично.

Если тот же текст изложен, например, на английском языке

ке, то может оказаться необходимым использование еще и прошлого опыта и самообучения, что для ЭВМ еще менее доступно. Наконец, необходимость принятия решения по этому тексту может вызвать в процессе творчества такой интенсивный обмен информацией между вербализованным и невербализованными уровнями мышления, что имитация процесса принятия решения может оказаться, что бывает довольно часто, не под силу ЭВМ.

Анализируя сходство и различие мышления и "машинного интеллекта", Ю.А. Шрейдер [13] справедливо замечает: "В науке хорошо известно, каким огромным достижением является разумная постановка проблемы или, наоборот, выяснение бессмысленности какого-то вопроса. И может быть, самая важная особенность человеческого разума - это способность отказаться от попыток достижения бессмысленной цели". Это замечание может оказаться справедливым и по отношению к проблеме создания эффективных АИПТ в силу сформулированного выше принципа.

Однако принцип предельных возможностей АИПТ дает не только негативный результат. С его помощью можно прогнозировать области наиболее эффективного применения АИПТ и, в частности, АИС.

Для этого необходимо перед началом работ по имитации проанализировать с помощью специалистов, главным образом, психологов конкретные творческие задачи, решаемые человеком. Результатом такого анализа должна быть приближенная оценка доли вклада того или иного способа обработки информации в общем объеме решаемой задачи. Если основным является формально-логический способ, то применение АИПТ может дать весьма эффективный результат, разумеется, если выбран математический аппарат, адекватный задаче.

Если основную роль играет, например, ассоциативное мышление, то применение ЭВМ в качестве технического средства АИПТ неэффективно и надо искать другой машинный способ обработки информации, например, [4].

Последовательное применение предложенного выше принципа предельных возможностей позволяет сделать вывод и о том, что ЭВМ не является единственной и, тем более не последней "инстанцией", к которой надо обращаться специалистам, работающими над проблемами АИПТ. Это понимал еще в начале второй половины века К. Шеннон [12]: "Эффективное машинное решение таких задач, как распознавание образов, автоматический перевод и т.д. может потребовать создания вычислительного уст-

ройства совершенно другого типа, чем то, которым мы располагаем сегодня. Как мне представляется, что будет устройство, естественное функционирование которого происходит при помощи образов, понятий и смутных подобий, а не с помощью последовательных операций над числами в десятичной системе счисления".

Чтобы наиболее рельефно подчеркнуть важную роль выбора путей совершенствования технических средств и способов обработки информации в лингвистике, сравним два возможных метода имитации, например, ассоциативно-образного мышления. Придя, наконец к пониманию роли образов, М. Минский [8], например, предложил систему фреймов для описания ситуаций. Однако, их структура такова, что операции над фреймами и их компонентами ведется у М. Минского алгоритмическим, т.е. формально-логическим способом обработки информации со всеми вытекающими из принципа предельных возможностей ограничения.

Между тем, при использовании, например, идеи голографического процессора [4] реализуются операции ассоциативной записи цепочек изображений (образов). Результатом работы процессора в силу этого являются неизвестные ранее (скрытые) пространственные и временные связи между элементами цепей и интерполяция образов. Значение такого принципа обработки информации трудно переоценить, поскольку выявление скрытых связей между понятиями с учетом непрерывности семантического поля весьма необходимо для решения задач понимания текстов и принятия решений с помощью АЯС.

Резюмируя вышеизложенное, нельзя не вспомнить слова академика А.Н. Колмогорова "В развитом сознании современного человека аппарат формального мышления не занимает центрального положения. Это скорее некое "вспомогательное вычислительное устройство", запускаемое в ход по мере надобности. Кибернетический анализ работы развитого человеческого сознания в его взаимодействии с подсознательной сферой еще не начат". Эти слова, сказанные еще в 1964 году, справедливы и по сей день.

Главная мысль настоящей работы состоит в том, что работа над проблемой АЯС, должна включать в себя следующие этапы: анализ лингвистами и психологами вклада различных способов обработки информации в решении человеком той или иной творческой задачи, определение разумных границ использования АЯС, выбор технических средств, адекватных основным способом об-

работки информации, с учетом изложенного выше принципа предельных возможностей АИПТ.

### Литература

1. Адамар Ж. Исследование психологии процесса изобретения в области математики. М., "Советское радио", 1974.
2. Алексакин В.И. Мышление и кибернетика, М., 1971.
3. Бассин Ф.В. Проблемы бессознательного. М., 1968.
4. Быховский В.К. Принципы построения ассоциативного оптического процессора. Проблемы голографии, вып. 2, изд. МИРЭА, 1973.
5. Брушлинский А.В. Мышление и прогнозирование. М., "Мысль", 1979.
6. Дмитриев А.И. Взаимосвязь осознанного и неосознанного в процессе отражения действительности. - В кн.: Современные проблемы теории познания диалектического материализма. М., 1970.
7. Мильнер Б.З. Проблемы управления в современной Америке. М., "Знание", 1974.
8. Минский М. Фреймы для представления знаний. М., "Энергия", 1979.
9. Морозов М.Н. Творческая активность сознания. Киев, 1976.
10. О сущности жизни. М., "Наука", 1964.
11. Психология мышления. "Прогресс", 1965.
12. Шеннон К. Работы по теории информации и кибернетике. М., 1963.
13. Шрейдер Ю.А. Присущ ли машине разум? - "Вопросы философии", 1975, № 2.



SOME PROBLEMS CONNECTED WITH THE CONSTRUCTION OF  
AUTOMATIC IMITATORS OF CREATIVE PROCESSES

V.Safonov

S u m m a r y

The reading and construction of texts representing the results of creative processes are considered to be of immense importance. This also accounts for the rise and development of natural language based automatic interaction systems. In his paper the author examines the interaction systems of the type that lack a strictly formalized structure of text buildup (e. g. texts on statistics, economics, etc.). He tries to find an answer to the question of what causes the difficulties in artificial intelligence investigations and, on the other hand, to draw the attention of automatic interaction system experts to the methods of analysing these difficulties and to the existence of ways of overcoming them at the present stage.

## ЗНАЧЕНИЕ, СМЫСЛ И МЫСЛЬ

И. Сильдин

Создание систем искусственного интеллекта, оперирующих естественным языком, выдвигает на первый план проблему моделирования языка не как самостоятельной лингвистической системы, а как средства познания. Это предполагает понимание познавательной сущности языка и его функционирования /1/.

х х  
х

(I) Чувственное и языковое познание организованы у человека в принципе одинаково. Познавательный процесс включает два аспекта. Один из них представляет узнавание объектов познания, другой - понимание связей между этими объектами. Эти познавательные связи проявляются как роли участников связи. Роль выражает функцию участника связи. Таким образом, обеспечивается единство структуры и ее функциональной стороны в познании реальности.

Чувственное познание всегда имеет дело с конкретными объектами. Объективной основой узнавания этих объектов является их сходство в определенных границах, т.е. единство единичного и общего. На этом сходстве основывается и субъективное узнавание этих внешних объектов.

Объективной основой познания ролей объектов является сходство связей между ними в реальности. В случае физических объектов, например, таковыми являются связи между ЧАСТЬЮ и ЦЕЛЫМ, ПРИЗНАКОМ и ЦЕЛЫМ, между ЧАСТЬЮ ЧАСТИ и ЧАСТЬЮ, между ПРИЗНАКОМ ПРИЗНАКА и ПРИЗНАКОМ и т.д. Узнавая ЧАСТЬ некоторого объекта, мы знаем, что играет роль ЦЕЛОГО, узнавая ЦЕЛОЕ, мы знаем, что выступает в роли его ЧАСТЕЙ, ПРИЗНАКОВ и т.д.

В пространственных ситуациях, движениях, событиях, между участвующими объектами имеются свои характерные связи, участники которых имеют соответственные функциональные роли. В случае чувственного познания все эти связи находятся, так сказать, при объектах, мы видим объекты в их ролях.

Психологически эти связи обнаруживаются через различные ассоциации (предметные, пространственные, временные).

Язык и его применение построены на тех же принципах, что и чувственное познание.

Слова в отдельности обозначают те объекты познания, которые человек научился узнавать и которые он считает необходимым отделить от других объектов при помощи слова (собака, кошка, нога, красный, лес, бег и т.д.). Знания о самих этих объектах составляют значение соответствующих слов. Значение слова состоит из познавательных категорий, которые характеризуют соответствующий объект познания. Это особая проблема, которую мы в данной работе не будем рассматривать подробно.

При составлении предложения мы связываем слова между собой в ту же ролевую структуру, в какой находятся данные объекты познания в реальности. Так формируется познавательная, или гносеологическая структура предложения.

В предложении одни слова выступают в роли ЦЕЛОГО, другие - в роли ЧАСТЕЙ этого ЦЕЛОГО, третьи - в роли ПРИЗНАКОВ, четвертые - в роли ДВИЖЕНИЯ, пятые являются ПРИЗНАКАМИ ПРИЗНАКОВ и т.д. Эти слова в ролях образуют описания представленных в предложении объектов. Такие описания объектов мы называем ДАТУМАМИ. Датумы находятся, в свою очередь, в тех или иных связях, в зависимости от того, что описывается предложением. Отсюда исходят роли датумов, которые выражают их функцию в описываемой ситуации или событии.

Таким образом, слово в каждом конкретном употреблении, в дополнение к своему значению, имеет еще и роль, иначе говоря, смысл, его употребления. Смысл датума определяется его ролью в предложении. А вся познавательная структура предложения в целом образует мысль, которую мы хотим выразить этим предложением.

Именно эта ролевая структура текста и делает очень гибким употребление слов при выражении различных мыслей. Одно и то же слово может выступать в разных предложениях в различных ролях, наряду с другими разными ролями и их исполнителями. Слово участвует в выражении различных мыслей, приобретая, таким образом, многочисленные ролевые связи с другими словами. Эти связи и составляют наши знания об объектах, ситуациях и событиях окружающего мира.

По такой же ролевой структуре построены тексты и происходит их понимание. А в тексте уже целые предложения выступают в тех или иных ролях в отношении ситуаций или событий,

там описываемых. И эта ролевая (познавательная) структура выражает мысль соответствующего текста.

Вплоть до уровня предложения морфологические и синтаксические признаки языка являются руководящими при образовании познавательной структуры. Это позволяет нам образовывать как значения самих слов, так и знания, вытекающие из связи между словами, т.е. смысл. Это позволяет также осуществлять накопление знаний без предварительных семантических знаний. Если бы это происходило иначе, то знания у человека должны были бы предшествовать их образованию и накоплению при помощи языка.

Но при понимании текстов и при их составлении человек одновременно оперирует как своими знаниями, так и грамматической формой слов и синтаксической формой предложения, которые в значительной мере дублируют друг друга. И только при возникновении противоречия между новыми и старыми знаниями возникает дилемма: чему верить - своим старым знаниям или языковой форме сказанного.

Развитие человеческих знаний филогенетически и онтогенетически также характеризуется двумя процессами. Во-первых, усвоением новых познаваемых объектов, во-вторых, все более глубоким и детальным пониманием познавательных связей между этими объектами, т.е. развитием ролевой структуры.

Создание кибернетики в наши дни выражает развитие познавательной структуры. Например, понимание и учет обратной связи ролей сторон этой связи (понимание того, что ДЕЙСТВУЮЩИЙ является одновременно ПРИНИМАЮЩИМ обратного ДЕЙСТВИЯ).

Нам кажется, что многие знаковые системы так же, как и естественный язык, имеют две стороны: знаки-объекты и межнаковые связи, при которых знаки выполняют разные роли, функции, имеют различный смысл. Примером этого может служить арифметика со своими знаками-числами и связями-действиями, или алгебра со своими буквами и действиями и т.д. И в математике, как знаковой системе, развитие познания выражается в открытии новых знаков и их связей, а также в выражении при их помощи все более сложных математических познавательных структур.

(2) Знания людей (память) построены на основе познания внешнего мира - то есть по принципу той же познавательной структуры, а не как языковые - семантические - структуры.

Уже сто лет тому назад известный русский физиолог

Н.М. Сеченов утверждал: "Всякий внешний предмет или явление фиксируется в памяти и воспроизводится в сознании в трех главных направлениях: как член пространственной группы, как член преемственного ряда и как член сходственного ряда" /2/.

Организацию наших знаний в виде познавательных структур по принципу объектов, пространственных групп и временных рядов подтверждают исследования советского психолога А.Р. Лурия о построении и топография памяти /3/.

В памяти четко отделены значения слов от знания о пространственных и временных познавательных структурах.

Знания о значениях самих слов, обозначающих объекты познания, организованы на основе связей: ЧАСТЬ-ПРИЗНАК-ЦЕЛОЕ. Рост человеческих знаний добавил к этому еще родо-видовые связи, стороны которых выполняют соответствующие роли. Вместе они образуют вертикальную (логическую) систему связей между словами.

Знания, образованные при помощи слов о больших познавательных межобъектных структурах, также играют соответственные роли. Наши знания в том и заключаются, что мы знаем, что находится (или не находится) в той или иной роли, если известны те или другие исполнители остальных ролей пространственной ситуации или события.

Такая ролевая структура знаний позволяет сохранить в памяти мысли, независимо от их первоначальной языковой формы. И с другой стороны, знание слов не гарантирует еще возможности выражения мысли. В основе сохранения мысли лежит сохранение соответствующей ролевой структуры, которая заполняется при забывании первоначальных слов другими, пригодными для этих ролей словами, близкими по значению к первоначальным словам.

На наш взгляд, ролевой способ представления знаний является именно тем, который Д.А.Поспелов называет "языком смыслов" и "модельным языком", считая их основой создания систем И.И. /4/.

(3) Из такой организации знаний в памяти следует два аспекта человеческого мышления: формально-логическое мышление и ролевое мышление, их взаимодействие и единство.

Организация знаний по вертикальным связям позволяет сделать формально-логические выводы как внутри одного вида (понятия), так и между родо-видовыми понятиями. Силлогизмы можно построить на основе всех признаков понятия (значения) и

получить соответствующие формально-логические выводы. Какие выводы позволяет получить формально-логическое мышление, этого вопроса мы здесь ввиду общеизвестности предмета не касаемся.

Рольное мышление раскрывает пространственные, временные и связанные с движением связи между объектами. Суть ролевого мышления состоит в том, что исходя из какого-то объекта, выступающего в какой-то роли, мы знаем на основании более ранних связей, кто или что может выступать в других ролях познавательной структуры.

Приведем только несколько примеров о возможностях ролевого мышления.

а) Все языковые вопросы относятся к ролям и направлены на выяснение той или другой роли. Понимание языковой формы вопросов не являются трудным, поскольку вопросы имеют такую же познавательную структуру, как наши знания. В вопросе дана часть заполнителей ролей и необходимо найти заполнитель той роли, на которую направлен вопрос (кто, что, с кем, почему и т.д.). Нетрудно найти требуемый ответ среди наших ролевых знаний, если это знание там явно, эксплицитно имеется.

Положение осложняется в том случае, когда вопрос задан словами, отличающимися от тех, которые были использованы при вводе знания. Тогда возникает проблема замены слов.

б) Замена слов связана со сходством обозначаемых ими явлений. Для человека эта проблема разрешена тем, что близкие по значению слова находятся в памяти поблизости. В системе И.И. эту задачу можно будет решить с помощью применения вышеупомянутого вертикально-логического принципа организации знаний.

в) Обобщение. Знания поступают в память относительно различных множеств, от одного до множества "все". Поскольку знания организованы в памяти в виде единства отдельного и общего, обобщение представляет собой установление наличия какого-то аспекта сходства при том или ином множестве.

д) Цепи умозаключений. При них возникают два вопроса: какие заключения человек делает и каким образом человек знает, какими знаниями в том или другом случае он должен воспользоваться?

Что касается того, какие заключения человек делает, то на это можно ответить так: он старается заполнить пробелы в познавательной структуре знаний, имеющих существенное значе-

ние для понимания текста, для ответа на вопрос, для выяснения интересующих его аспектов и т.д.

Выбор знаний, которыми нужно пользоваться при выводах, следует также из познавательной структуры. Повседневное мышление занимается преимущественно поиском носителей отсутствующих ролей частично данной познавательной структуры или же увязыванием познавательных структур, в целях понимания недосказанного, но в тексте имплицитно имеющегося знания.

Если наши знания не имели бы такой познавательной-ролевой структуры, то отсутствовали бы структурально-функциональные основы для осуществления выбора из числа имеющихся знаний и возможных выводов.

Нам кажется, что эти трудности относятся ко всем системам, обоснованным на семантическом представлении языка.

Выводы, которые делаются, можно разделить на три уровня: (1) выводы на микроуровне о самих объектах, действиях, отношениях. Большое значение в данном случае имеет дополнение знаний, приобретаемых посредством языка, чувственно поступающими знаниями. Здесь проявляется совместная работа аналогично построенного чувственного и языкового познания.

(2) выводы на макроуровне делаются на основе познавательной структуры, лежащей в основе предложения. Здесь человек оперирует, главным образом, своими знаниями о пространственных группах и временных рядах.

(3) сверхуровнем является текст и сделанные из него выводы. Здесь в виде выводов происходит образование познавательных структур крупных ситуаций, событий, описанных при помощи предложений.

Процедуры переработки знаний при выводе заключений соединяют логическое и ролевое мышление и представляют собой твердые типы. Варьируется очередность пользования ими в зависимости от характера имеющихся знаний и преследуемой цели (вывода). Нам кажется, что эти процедуры похожи на разработку данных при системном анализе, с чем сравнивает мышление Г.С. Поспелов.

(4) Создание и сущность систем искусственного интеллекта самым тесным образом связаны со степенью разумности этих систем.

Познание не пассивный процесс, оно самым тесным образом связано с отношением системы к познаваемому окружению. Это отношение к познаваемому, или РАЗУМНОСТЬ, является движущей

и направляющей силой мышления.

Первоначально разумность имеет биологическо-потребительский характер. В развитии разумности человека решающее значение имеет процесс труда и социальный образ жизни.

Разумность добавляет к познавательной структуре внешнего мира еще одну часть - "Я". Так образуется когнитативная структура или структура мышления, которая соединяет в себе знания и отношения к ним, что и определяет использование этих знаний.

И в системах И.И. их разумность проявляется в отношении системы к поступающим знаниям. Эта пока проблема будущего. Но при этом возникают острые этические проблемы, касающиеся того, какое направление будет иметь разумность систем И.И. - творческое или *vice versa*. Что же представляет собой, например, творческая разумность в той или другой области науки? Об этом знания самих людей еще очень скромны. Но нам кажется, что эта проблема самым тесным образом связана с ролевым представлением знаний и процедурами их обработки.

Нет сомнения в том, что проблема создания разумных гворческих систем искусственного интеллекта в будущем будет разрешена. Тогда системы И.И. из систем, выполняющих заданные человеком задачи, т.е. использующих человеческую разумность, превратятся в разумные системы и станут действительно интеллектуальными, творческими помощниками человека.

#### Литература

1. Сильдмяз И. Гносеологическая структура текста и знаний. - Уч. зап. Тартуского гос. ун-та. Труды по искусственному интеллекту II. Тарту, 1980.
2. Сеченов И.М. Элементы мысли М.-Л., 1943, с. 138.
3. Лурия А.Р. Высшие корковые функции человека. М., 1962.
4. Поспелов Д.А. Диалоговые системы: трудности и перспективы. - Уч. зап. Тартуского гос. ун-та. Труды по искусственному интеллекту II. Тарту, 1980.



## MEANING, SENSE, THOUGHT

I.Sildmäe

### S u m m a r y

Creation of AI systems operating with natural language sets forward a problem of language modelling not as an independent linguistic system, but as a means of cognition. The cognitive essence of language manifests itself through the gnosiological structure of the text. Taken separately, words possess only meaning. In a text words stand in certain gnosiological roles. These roles add a definite sense to the meaning of a word according to the situation.

The gnosiological structure of a sentence (text) forms a thought which one expresses by this sentence (text).

## ДИАЛоговая СИСТЕМА ПРИНЯТИЯ РЕШЕНИЙ С КОНТРОЛЕМ СОСТОЯНИЯ ЛПР

Д.И. Шапиро, В.Н. Васильев

Существует широкий класс многопроблемных задач, в которых каждая из независимых проблем (объектов) характеризуется своими локальными критериями и ограничениями, а также набором признаков. На процедуры оценки и выбора влияют свойства человеку неопределенности, а возможности его (ресурсы) ограничены. Ситуационная информация в общем случае представляется как в численном виде, так и в виде текста.

Критерии и ограничения (как локальные, объектов  $F_i:Q_i$ , так общий -  $Q_0$ ) заданы в виде текста.

Важность отдельных признаков объектов  $x_i$  также, как и формулировки критериев  $Q_i$ , могут меняться в динамике в зависимости от конкретной ситуации или состояния ЛПР.

Пользователю удобно в своей повседневной деятельности выбирать решения в режиме диалога с ЭВМ.

Реализация диалоговой процедуры принятия решений в системах с дисплеем в качестве внешних устройств позволяет достаточно оперативно проанализировать возможные рекомендации (например, в период подготовки к заседанию или во время заседания). В подобной диалоговой процедуре рассматриваются разнообразные проблемы (мнение о типе продукта, приблизительно подходящего по общности, достаточно ли хороши альтернативы, найдется ли лучший путь решения, или нужно прекратить поиски, выделение одной из альтернатив как "лучшей" и ее проверка жизнью и т.д.), которые сводятся к двум типам вопросов:

- "что будет, если данный вариант довести до конца"?
- "что будет, если в динамике процесса меняется решение"?

Первый тип вопросов связан с необходимостью получения набора рекомендаций  $\{R_i\}$  для выбора одной из них ЛПР  $F_0$ .

Второй тип вопросов возникает в случаях, когда в динамике появляются новые факторы (изменилась важность объектов  $F_i$  или взаимная важность признаков  $x_i$ ).

Решение подобных задач, по-видимому, целесообразно осуществлять на основе ситуационной модели выбора рациональных рекомендаций (СМВРР).

В соответствии с функциями, реализуемыми в модели, она состоит из следующих макроблоков:

Макроблок "Представление" осуществляет формирование описания задачи и условий ее решения.

В этом макроблоке формируются: набор ситуационных данных

$$S_t = \{P, x_i, \underline{x}_i\}, \text{ где:}$$

$P$  - реквизиты;

$x_i$  - данные, представленные в количественной форме;

$\underline{x}_i$  - данные, представленные в качественной форме.

(При этом в зависимости от конкретной задачи данные представляются, например, в виде упорядоченного набора или в виде статического фрейма).

- Критерии и ограничения для объектов  $F_0$  и  $F_i$  (соответственно  $Q_0, Q_i, L_0, L_i$ ).

- Система отношений  $\Gamma_i$  между объектами  $F_0$  и  $F_i$  множества элементарных управлений  $u \in U$  для  $F_0$  и  $v_i \in V_i$  для  $F_i$  соответственно.

- Система оценочных шкал  $c^j \in C$

Элементы набора ситуационных данных, имеющие количественную форму, приводятся к виду

$$\bar{x}_i = \frac{x_i - \min x_i}{\max x_i - \min x_i}$$

(Блок П-1с).

Формальное представление качественных категорий, встречающихся в текстах, которыми описываются данные о состоянии и изменении системы и среды, критерии  $Q_0, Q_i$  и ограничения  $L_0, L_i$  осуществляется на основе соответствующей [6] методики формализации. При этом все расплывчатые категории выражаются своими функциями принадлежности  $\mu(x) \in \mathcal{F}(\cdot)$ , где  $\mathcal{F}(\cdot)$  - множество всех функций принадлежности.

В функции макроблока "Представление" входит организация диалогового режима между ЛПР (пользователь) и ЭВМ.

Запрос с помощью дисплея посредством переходного блока Пер. Б. подается на выход системы в виде текста на естественном языке и численных данных (которые, как отмечено выше, обработаны на блоке П-1).

Текст, проходя через лингвистический анализатор (Лин А) и блок П-2, преобразуется в "понятный ЭВМ" язык, в котором расплывчатые категории представлены соответствующими значениями функций принадлежности, хранящихся в блоке формализо-

ванных расплывчатых категорий (БФРК) главной базы данных.

В диалоговой процедуре используются возможности модифицированной системы ДИЛОС, разработанной В.М. Брябриным и Г.В. Сениным [1].

В макроблоке "Контроль" на основе упорядоченного формализованного набора ситуационных данных  $\bar{S}_t$  реализуются следующие процедуры:

Контроль (отслеживание) изменения во времени некоторых параметров  $x_i^j$  (блок К-1).

Несколькими блоками реализуется субъективное представление ЛПР об особенностях объектов  $F_i$ .

В блоке К-2 по запросу осуществляется контроль постоянства или изменения локальных критериев.

В случае изменения или уточнения критерия проводится анализ принадлежности нового критерия  $Q_i'$  к данному классу критериев  $Q_i \in \bar{Q}$ ;  $Q_i' \in \bar{Q}$  (например, для случая сжатия критерия или его ослабления) или к другому классу критериев  $Q_i \in \bar{Q}$ ;  $\bar{Q}_i' \in \bar{Q}$ ;  $\bar{Q} \cap \bar{Q}' \neq \emptyset$ .

В блоке К-3 по запросу производится контроль и анализ типичных свойств  $B_K$  (физических координат, мощности ресурсов) и, на основании близости подобных данных, осуществляется объединение  $B_K$  в соответствующие узкие группы.

В макроблоке "Оценка" осуществляются процедуры оценки объектов  $F_i$  по значению из признаков  $x_i^j$ . При этом в зависимости от конкретной задачи фиксируются значения по совокупности независимых шкал (блок О-1).

$$c^j: x_i^j \rightarrow y_i^j$$

(например, для некоторых экономических показателей) определяется рассогласование между состоянием одноименных признаков ЛПР  $F_0$  и объектов  $F_i$  по совокупности независимых шкал (блок О-2)

$$c^j: x_0^j - x_i^j \rightarrow y_i^j$$

(например, для динамических параметров - координаты, скорость, ускорение и т.д.) или определяется состояние объекта  $F_i$  по значениям координат  $y_i^j$  в пространстве оцененных параметров  $Y$ . При этом каждый объект  $F_i$  представлен вектором

$$y_i = [y_i^1, \dots, y_i^m]; y \in Y$$

На произвольных этапах многопроблемной многошаговой процедуры может возникнуть необходимость оценки изменившейся ситуации (на основе субъективных представлений ЛПР) в пространстве оцененных параметров.

Реализация подобных функций осуществляется соответствующими блоками, также входящими в макроблок "Оценка". Блок 0-4 реализует формальное представление взаимной важности признаков в виде текста ("  $y^1$  слабо важное  $y^2$ ", "Отношение  $y^2$  и  $y^3$  между менее важно и эквивалентно", "  $y^3$  слабо важнее, чем  $y^1$  ) с помощью оператора:

$$\alpha = \bar{\lambda} G.$$

где  $G$  - эрмитова матрица,  
 $\bar{\lambda}$  - собственный вектор.

Тогда каждый объект  $F_i$  в пространстве оцененных параметров представлен вектором  $\hat{y}_i = [(y_i^1) \alpha_{1j}, \dots, (y_i^m) \alpha_{mj}]$ . В блоке 0-5 по запросу осуществляется оценка системы предпочтения объектов  $F_i$  и соответствующих отношений  $\Gamma_i$  на основе качественных категорий. При этом реализуется алгоритм

$$((y^k) \alpha_k \Gamma_j) \Psi((y^l) \alpha_l \Gamma_j); j \neq k$$

$\Psi$  имеет градации: индифферентно - слабое предпочтение-сильное предпочтение - очень сильное предпочтение.

В макроблоке "Выбор" осуществляется определение стратегии (рациональных рекомендаций) ЛПР в многопроблемной многошаговой процедуре.

В зависимости от конкретной задачи рассматривается план применения методов или способ поведения (состоящий из последовательности элементарных управлений), которые обеспечивают рациональное (в смысле критериев и ограничений) действие ЛПР во всех  $\Gamma_i$  на каждом  $t$ .

Формальным аппаратом для реализации процедуры выбора является описанная в /5/ расплывчатая интегральная игра.

Таким образом, ситуационная модель выбора рациональных рекомендаций, основанная на расплывчатой интегральной игре, позволяет варьировать в диалоговом решении элемента набора ситуационных данных  $\{x_i^j\}$ , сравнительную важность отдельных параметров оценки  $(y_i^j) \alpha_{ij}$ , важность конкретных объектов  $(\Psi_i)$ , критерии  $Q_0$  и  $Q_i$ , а также ограничение  $L_0$  и  $L_i$ . Выработка рациональных рекомендаций ЛПР связана с его состоянием.

Использование СМВРР, алгоритмированной на ЭВМ, позволяет облегчить деятельность ЛПР как человека-оператора, так и организационного руководителя.

Однако утомление ЛПР или его реакция на стресс могут привести к нарушениям процесса принятия решений.

Коррекция действия ЛПР осуществляется на основе контроля состояния некоторых психофизиологических параметров, характеризующих эмоциональную стоимость работы, в частности, симпатно-адреналовой системы (САС).

На основе изучения САС удалось создать шкалу эмоциональной напряженности работников ряда операторских профессий, например, диспетчеров железнодорожного транспорта, телеграфистов, телетайпистов, а так же пользователей ЭВМ в организационных системах управления, находящихся под воздействием нервно-эмоционального стресса /2-4/.

Шкала является 9-ти позиционной, причем каждый интервал (позиция) представлен парой (номер, функция принадлежности).

Выявлены три типа ответа САС на стресс: гиперреакция гипореакция и отставленная (постстрессовая) реакция.

Гиперреакция характеризуется увеличенной активностью САС с неэкономным расходом ее гормонов и медиаторов и снижением ее резервных возможностей. Подобное состояние говорит о нарушении адаптации ЛПР к данным условиям труда и допустимо лишь в течение ограниченного времени.

Гипореакция характеризуется сниженной активностью САС, что сопровождается снижением выделения нейrogормонов. Подобное состояние связано с необходимостью учета этих данных при профотборе.

Пост-стрессовая реакция характеризуется увеличением активности САС не во время действия стресса, а в период отдыха ЛПР, что может значительно нарушить процессы их восстановления.

Анализ поведения ЛПР при рассмотренных типах реакций САС показал, что гиперреакция приводит к возбуждению, гипореакция - к торможению и пост-стрессовая - к возбуждению (степень возбуждения или торможения определяется на основе использования описанной выше шкалы)

Использование подобной шкалы позволило оценить эмоциональную стоимость труда конкретных ЛПР, определять период или степень утомления, прогнозировать периоды их восстановления, а также возможную степень отклонений действия ЛПР от

нормы в будущем. Кроме того, с помощью подобной шкалы можно оценивать работу и состояние не только клинически здоровых ЛПР, но и таких, которые находятся в ранней стадии развития гипертонической или язвенной болезни.

Функции контроля состояния ЛПР выполняются специальным блоком модели.

На основании результатов анализов и прогнозирования состояния ЛПР в этом блоке формируется (для самого ЛПР и системы внешнего контроля) оценки надежности его дальнейшей работы.

Кроме того, в этом блоке осуществляется контроль поведения ЛПР на качественной плоскости оценок "улучшение-ухудшение, быстро-медленно". Траектория на этой плоскости фиксирует последовательную процедуру выработки решений, временные параметры и т.д. и позволяет определить закономерности в выработке решений конкретных ЛПР, узкопрофессиональной группой, а так же отклонения от подобных закономерностей /5, 6/.

#### Литература

1. Брябрин В.М., Сенин Г.А. Диалоговая информационно-логическая система ДИЛОС. ВУАН СССР, 1977.
2. Васильев В.Н. Гигиена труда и профзаболевания, 1974, № 8.
3. Васильев В.Н. - В кн.: Стресс и его патогенетические механизмы. Кишинев, 1973, с. 25.
4. Васильев В.Н., Горшков В.А. - В кн.: Неспецифические заболевания в связи с профессиональными факторами. М., 1976, с. 45.
5. Шапиро Д.И. Расплывчатые интегральные игры. - В кн.: Модели и системы принятия решений. РПИ. Рига, 1979.
6. Шапиро Д.И., Лесонтьев А.А. Расплывчатые категории в проблеме принятия решений. - В кн.: Общете методические и прогностические проблемы. ИЯ АН СССР, 1978.

## A DIALOG SYSTEM FOR DECISION MAKING WITH FOLLOWING CHECK-UP

D. Shapiro, V. Vasilyev

### S u m m a r y

There exists a wide class of multiproblem tasks where every independent problem (object) can be characterized by its own local features and restrictions. The procedures of evaluation and choice by man are influenced by uncertainty and restricted resources.

The implementation of a dialog procedure of decision making allows us to analyze the possible recommendations quite effectively. This procedure makes it possible to consider various problems which can be reduced to two types of questions:

- What happens if the given version is carried out up to the end?

- What happens if the decision is changed during the process?

The adjustment of the activities of an operator is performed on the basis of a check-up of the state of some psychological and physiological parameters characterising the emotional value of the work.



Ученые записки  
Тартуского государственного университета.  
Выпуск 551.  
ЛОГИКО-СЕМАНТИЧЕСКИЕ ВОПРОСЫ  
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА.  
Труды по искусственному интеллекту III.  
На русском языке.  
Резюме на английском языке.  
Тартуский государственный университет.  
ЭССР, 202 400, г. Тарту, ул. Пялсона, 14.  
Ответственный редактор А. Бим.  
Корректоры Н. Чикалова, М. Садумев.  
Подписано к печати 05.12.1980.  
МВ 11703.  
Формат 30x45/4.  
Бумага писчая.  
Машинопись. Ротапринт.  
Учетно-издательских листов 7,82.  
Печатных листов 8,5.  
Тираж 500.  
Заказ № 1344.  
Цена 1 руб. 20 коп.  
Типография ТТУ, ЭССР, 202400, Тарту, ул. Пялсона, 14.